Quantum Leap
From Dirac and Feynman, Across the Universe, to Human Body and Mind

Vladimir G Ivancevic
Tijana T Ivancevic

World Scientific
Quantum Leap

From Dirac and Feynman,
Across the Universe,
to Human Body and Mind
This page intentionally left blank
Quantum Leap
From Dirac and Feynman,
Across the Universe,
to Human Body and Mind

Vladimir G Ivancevic
Defence Science & Technology Organisation, Australia

Tijana T Ivancevic
University of South Australia, Australia
Dedicated to:

Nick, Atma and Kali
Preface

*Quantum Leap: From Dirac and Feynman, Across the Universe, to Human Body and Mind* is a monographic textbook in the field of quantum physics and its relations to human mind and body. The book has seven Chapters. The introductory Chapter gives both non–mathematical and mathematical preliminaries for understanding the text. The second Chapter presents the basics of both non–relativistic and relativistic quantum mechanics. The third Chapter introduces Feynman path integrals and their application to quantum fields and string theory, as well as some non–quantum applications. The fourth Chapter presents quantum universe in the form of loop quantum gravity and quantum cosmology. The fifth Chapter turns to the human body, applying quantum mechanics and electrodynamics to electro–muscular stimulation. The sixth Chapter presents modern quantum games and quantum computers. The seventh Chapter develops quantum theory of the mind and consciousness.

This book can be used as a textbook of a two–semester course in quantum physics and its modern applications, at the graduate (or, higher undergraduate) level. Some parts of the book can be also used by engineers, biologists, psychologists and computer scientists, as well as applied mathematicians, both in industry and academia. It includes a comprehensive bibliography on the subject and a detailed index.

Adelaide, V. Ivancevic, Defence Science & Technology Organisation, Nov, 2007 Australia, e-mail: Vladimir.Ivancevic@dsto.defence.gov.au

T. Ivancevic, School of Mathematics, University of South Australia e-mail: Tijana.Ivancevic@unisa.edu.au
Acknowledgments

The authors wish to thank Land Operations Division, Defence Science & Technology Organisation, Australia, for the support in developing the Human Biodynamics Engine (HBE) and all HBE-related text in this monograph and Knowledge-Based Intelligent Engineering Systems Centre (KES), University of South Australia.

We also express our gratitude to the World Scientific Publishing Company, and especially to the Editors, Ms. Zhang Ji and Ms. Lai Fun Kwong. Finally, the book approval from Dr. K. K. Phua, Chairman and Editor-in-Chief of World Scientific is gratefully acknowledged.
# Contents

*Preface* vii

1. Introduction 1
   1.1 Soft Introduction to Quantum Mechanics . . . . . . . . . . 1
   1.2 Hilbert Space . . . . . . . . . . . . . . . . . . . . . . . . 7
      1.2.1 Quantum Hilbert Space . . . . . . . . . . . . . . . . 7
      1.2.2 Formal Hilbert Space . . . . . . . . . . . . . . . . . 8
   1.3 Human Intelligence, Mind and Reason . . . . . . . . . . . . 11
      1.3.0.1 Human Reason . . . . . . . . . . . . . . . . 13

2. Elements of Quantum Mechanics 19
   2.1 Basics of Non–Relativistic Quantum Mechanics . . . . . . . 19
      2.1.1 Canonical Quantization . . . . . . . . . . . . . . . . 23
      2.1.2 Quantum States and Operators . . . . . . . . . . . . 24
      2.1.3 Quantum Pictures . . . . . . . . . . . . . . . . . . . . 30
      2.1.4 Spectrum of a Quantum Operator . . . . . . . . . . . 32
      2.1.5 General Representation Model . . . . . . . . . . . . . 35
      2.1.6 Direct Product Space . . . . . . . . . . . . . . . . . . 36
      2.1.7 State–Space for n Quantum Particles . . . . . . . . . . 37
   2.2 Relativistic Quantum Mechanics and Electrodynamics . . . . 40
      2.2.1 Difficulties of the Relativistic Quantum Mechanics . . . 40
      2.2.2 Particles of Half–Odd Integral Spin . . . . . . . . . . 43
      2.2.3 Particles of Integral Spin . . . . . . . . . . . . . . . . 45
      2.2.4 Dirac’s Electrodynamics Action Principle . . . . . . . . 54
      2.2.5 Dirac Equation and Formal QED in Brief . . . . . . . . 60
      2.2.6 Lorentzian Space-Time and Gravity . . . . . . . . . . . 62
3.4.5 Topological Phase Transitions and Hamiltonian Chaos
  3.4.5.1 Phase Transitions in Hamiltonian Systems . 253
  3.4.5.2 Geometry of the Largest Lyapunov Exponent 256
  3.4.5.3 Euler Characteristics of Hamiltonian Systems 260

4. Quantum Universe
  4.1 Search for Quantum Gravity . . . . . . . . . . . . . . . . . 265
    4.1.1 What Is Quantum Gravity? . . . . . . . . . . . . . 265
    4.1.2 Main Approaches to Quantum Gravity . . . . . . . . 266
      4.1.2.1 String Theory . . . . . . . . . . . . . . . . 266
      4.1.2.2 Loop Quantum Gravity . . . . . . . . . . . . 269
    4.1.3 Traditional Approaches to Quantum Gravity . . . . 273
      4.1.3.1 Discrete Approaches . . . . . . . . . . . . . 273
      4.1.3.2 Hawking’s Euclidean Quantum Gravity . . . 275
      4.1.3.3 Effective Perturbative Quantum Gravity . . 276
      4.1.3.4 QFT in Curved Space–Time . . . . . . . . . 276
      4.1.3.5 New Approaches to Quantum Gravity . . . 277
      4.1.3.6 Hawking’s Black Hole Entropy . . . . . . . . 283
  4.2 Loop Quantum Gravity . . . . . . . . . . . . . . . . . . . . 284
    4.2.1 Introduction to Loop Quantum Gravity . . . . . . . 284
    4.2.2 Formalism of Loop Quantum Gravity . . . . . . . . 291
    4.2.3 Loop Algebra . . . . . . . . . . . . . . . . . . . . . 292
    4.2.4 Loop Quantum Gravity . . . . . . . . . . . . . . . . 293
    4.2.5 Loop States and Spin Network States . . . . . . . . 295
    4.2.6 Diagrammatic Representation of the States . . . 298
    4.2.7 Quantum Operators . . . . . . . . . . . . . . . . . . 299
    4.2.8 Loop v.s. Connection Representation . . . . . . . . 300
  4.3 Quantum Cosmology . . . . . . . . . . . . . . . . . . . . . . 301
    4.3.1 Hawking’s Cosmology in ‘Plain English’ . . . . . . 302
    4.3.2 Theories of Everything, Anthropic Principle and
      Wave Function of the Universe . . . . . . . . . . . . . 308
      4.3.2.1 Quantum Mechanics . . . . . . . . . . . . . 311
      4.3.2.2 Scientific Reduction . . . . . . . . . . . . . 312
      4.3.2.3 Anthropic Reasoning in Quantum Cosmology 313
      4.3.2.4 Quantum State of the Universe . . . . . . 320
      4.3.2.5 From Quantum Mechanics to Quantum
      Gravity . . . . . . . . . . . . . . . . . . . . 325
    4.3.3 Quantum Gravity and Black Holes . . . . . . . . . . 330
    4.3.4 Generalized Quantum Mechanics . . . . . . . . . . . 348
Contents

4.3.4.1 Quantum Mechanics Today . . . . . . . . . 351
4.3.4.2 Spacetime and Quantum Theory . . . . . . 352
4.3.4.3 The Quantum Mechanics of Closed Systems 353
4.3.4.4 Quantum Theory in 3+1 Form . . . . . . . 356
4.3.4.5 Generalized Quantum Theory . . . . . . . 358
4.3.4.6 A Quantum Theory of Spacetime Geometry 359
4.3.4.7 Beyond Spacetime . . . . . . . . . . . . . . . 362
4.3.4.8 Emergence/Excess Baggage . . . . . . . . . 363
4.3.4.9 Emergence of Signature . . . . . . . . . . . . 364
4.3.4.10 Beyond Quantum Theory . . . . . . . . . . 366

4.3.5 Anthropic String Landscape . . . . . . . . . . . . . 367

4.3.5.1 The Landscape . . . . . . . . . . . . . . . . 373
4.3.5.2 The Trouble with de Sitter Space . . . . . . 378
4.3.5.3 De Sitter Space is Unstable . . . . . . . . . 381
4.3.5.4 Bubble Cosmology . . . . . . . . . . . . . . . 382
4.3.5.5 Cosmology as a Resonance . . . . . . . . . . 384

4.3.6 Top–Down Cosmology . . . . . . . . . . . . . . . . 388

4.3.6.1 Trace Anomaly Driven Inflation . . . . . . . 391
4.3.6.2 Instability of Anomaly–Induced Inflation . . 401
4.3.6.3 Origin of Inflation . . . . . . . . . . . . . . . 405

4.3.7 Cosmology in the String Landscape . . . . . . . . . . 409

4.3.7.1 Quantum State . . . . . . . . . . . . . . . . 411
4.3.7.2 Prediction in Quantum Cosmology . . . . . . 413
4.3.7.3 Anthropic Reasoning . . . . . . . . . . . . . . 417
4.3.7.4 Models of Inflation . . . . . . . . . . . . . . . 418
4.3.7.5 Prediction in a Potential Landscape . . . . . . 421
4.3.7.6 Alternative Proposals . . . . . . . . . . . . . . 422
4.3.7.7 Interpretations . . . . . . . . . . . . . . . . 423

4.3.8 Brane Cosmology . . . . . . . . . . . . . . . . . . . . 425

4.3.8.1 The Randall–Sundrum Brane World . . . . 428
4.3.8.2 Including a Bulk Scalar Field . . . . . . . . 441
4.3.8.3 Moving Branes in a Static Bulk . . . . . . . 451
4.3.8.4 Cosmology of a Two–Brane System . . . . . 456
4.3.8.5 Brane Collision . . . . . . . . . . . . . . . . 462
4.3.8.6 Open Questions . . . . . . . . . . . . . . . . 468

4.3.9 Hawking’s Brane New World . . . . . . . . . . . . 469

4.3.9.1 RS Scenario from AdS/CFT . . . . . . . . . 473
4.3.9.2 CFT on the Domain Wall . . . . . . . . . . . 478
4.3.9.3 The Total Action . . . . . . . . . . . . . . . . 484
5. Quantum Biophysics in Human Body: Electro–Muscular Stimulation

5.1 Basics of Electrical Muscular Stimulation

5.2 EMS Functor

5.2.1 Global macro–level of $EM\!S_{total}$

5.2.2 Local Micro–Level of $EM\!S_{total}$

5.2.3 Micro–Level Adaptation and Muscular Training

5.3 Electrical Stimulation Fields: $EM\!S_{fields}$

5.3.1 External Smooth Maxwell Electrodynamics

5.3.2 Internal Cellular Bio–Quantum Electrodynamics

5.4 Stimulated Muscular Contraction Paths: $EM\!S_{paths}$

5.4.1 External Anatomical Muscular Mechanics

5.4.2 Internal Myofibrillar Bio–Quantum Mechanics

5.5 Anatomical Geometry of the Face & Body Shape: $EM\!S_{geom}$

5.5.1 External Face & Body Geometry

5.5.1.1 Facial Curvatures and Their Deviations

5.5.1.2 Facial Riemannian Patch–Manifold

5.5.2 Cellular Muscle–Fat Geometry

6. Quantum Games and Quantum Computers

6.1 Nash’s Classical Game Theory

6.1.1 Basics of Classical Game Theory

6.1.2 Nash Equilibrium: The Focal Point of Game Theory

6.1.2.1 Background to Nash Equilibrium

6.1.2.2 The Rationalistic Interpretation

6.1.2.3 The Mass–Action Interpretation

6.2 Quantum Games

6.2.1 Quantum Strategies

6.2.2 Quantum Games

6.2.3 Two–Qubit Quantum Games

6.2.3.1 Prisoners’ Dilemma

6.2.3.2 Chicken Game

6.2.4 Quantum Cryptography and Gambling

6.3 Formal Quantum Games
6.3.1 Formal Framework ........................................ 572
6.3.2 Various Types of Formal Games ....................... 573
  6.3.2.1 Strategic Games .................................. 574
  6.3.2.2 Quantum Simultaneous Games .................... 574
  6.3.2.3 Quantum Sequential Games ....................... 576
6.3.3 Equivalence of Formal Games .......................... 577
6.3.4 Game Classes ............................................ 582
6.3.5 Summary on Formal Quantum Games .................. 582
6.4 Quantum Information and Computing .................... 585
  6.4.1 Entanglement, Teleportation and Information ....... 585
  6.4.2 Quantum Computing ................................ 588
    6.4.2.1 Qubits and Quantum Parallelism .............. 591
    6.4.2.2 Quantum Computers ............................. 591
    6.4.2.3 The Circuit Model for Quantum Computers ... 592
    6.4.2.4 Elementary Quantum Algorithms .............. 594
6.5 The Hardware for Quantum Computers ................... 595
  6.5.1 Josephson Effect and Pendulum Analog .............. 599
    6.5.1.1 Josephson Effect ................................ 599
    6.5.1.2 Pendulum Analog ................................ 600
  6.5.2 Dissipative Josephson Junction ..................... 602
    6.5.2.1 Junction Hamiltonian and its Eigenstates .... 604
    6.5.2.2 Transition Rate .................................. 606
  6.5.3 Josephson Junction Ladder (J JL) .................... 608
    6.5.3.1 Underdamped J JL ............................... 614
  6.5.4 Synchronization in Arrays of Josephson Junctions 619
    6.5.4.1 Phase Model for Underdamped J JL ............ 622
    6.5.4.2 Comparison of LKM2 and RCSJ Models ......... 628
    6.5.4.3 ‘Small–World’ Connections in J JL Arrays .... 630
7. Quantum Mind ............................................... 635
  7.1 Quantum Theory of Consciousness ...................... 635
    7.1.1 Human Mind ........................................ 635
    7.1.1.1 The Mind–Body Problem ......................... 640
    7.1.2 EPR Paradox and Bell’s Theorem ................. 655
    7.1.2.1 EPR Paradox ..................................... 655
    7.1.2.2 Bell’s Theorem .................................. 662
    7.1.3 Orchestrated Reduction and Penrose Paradox ...... 667
    7.1.3.1 Orchestrated Objective Reduction .............. 669
    7.1.3.2 The Orch OR Model ............................... 671
Quantum Leap

7.1.4 Physical Science and Consciousness . . . . . . . . . . 673
  7.1.4.1 Reducing the Consciousness to Its Elements . . . . . . 674
  7.1.4.2 Manipulation of Consciousness . . . . . . . . . . 682
  7.1.4.3 A Need for Paradigm–Shift . . . . . . . . . . 685
  7.1.4.4 Quantum Synapses . . . . . . . . . . . . . . 686
  7.1.4.5 Quantum Timing of Conscious Experience . . . . . 687

7.1.5 Quantum Brain . . . . . . . . . . . . . . . . . . . . . 687
  7.1.5.1 Biochemistry of Microtubules . . . . . . . . . . 687
  7.1.5.2 Kink Soliton Model of MT–Dynamics . . . . . . . 689
  7.1.5.3 Liouville–Hamiltonian Neurodynamics . . . . . . 692
  7.1.5.4 Dissipative Quantum Brain Model . . . . . . . . 698
  7.1.5.5 QED Brain . . . . . . . . . . . . . . . . . . 706

7.1.6 Theory of Mind and Matter . . . . . . . . . . . . . . . . 714
  7.1.6.1 Matter . . . . . . . . . . . . . . . . . . . . . 715
  7.1.6.2 Mind . . . . . . . . . . . . . . . . . . . . . . 721
  7.1.6.3 Unification . . . . . . . . . . . . . . . . . . . 726

7.1.7 Quantum Consciousness . . . . . . . . . . . . . . . . . 729
  7.1.7.1 The Experiment . . . . . . . . . . . . . . . . . 733
  7.1.7.2 Bio–Active Peptides . . . . . . . . . . . . . . . 736
  7.1.7.3 Drugs . . . . . . . . . . . . . . . . . . . . . . 737
  7.1.7.4 Evolutionary Advantage . . . . . . . . . . . . 738

7.1.8 Quantum–Like Psychodynamics . . . . . . . . . . . . . 738
  7.1.8.1 Life Space Foam . . . . . . . . . . . . . . . . . 739

7.1.9 Joint Action Psychodynamics . . . . . . . . . . . . . . 755

7.1.10 Feynman–Finsler Approach to Adaptation . . . . . . . . 759

Bibliography 763

Index 827
Chapter 1

Introduction

This chapter gives the preliminaries for comprehensive reading of the book: an intuitive introduction to quantum mechanics, a minimum necessary mathematics (of the Hilbert space), as well as some very basic facts about the human mind.

While appreciating the outstanding pioneering efforts of Werner Heisenberg and Erwin Schrödinger (and others), this book is mainly focused on quantum theory of Paul Dirac and Richard Feynman (with exception of the chapter four) and its applications to human mind and body. Note that Einstein’s summation convention over repeated indices is assumed in the whole text. To make it more readable for physicists, the book is intentionally less rigorous than in similar texts on graduate-level mathematical physics. Also, every section begins more intuitively, and later develops more formally. For all additional explanations regarding the used mathematical formalisms, the reader is referred to our book Applied Differential Geometry, World Scientific, 2007 [Ivancevic and Ivancevic (2007b)].

1.1 Soft Introduction to Quantum Mechanics

According to quantum mechanics, light consists of particles called photons, and the Figure 1.1 shows a photon source which we assume emits photons one at a time. There are two slits, A and B, and a screen behind them. The photons arrive at the screen as individual events, where they are detected separately, just as if they were ordinary particles. The curious quantum behavior arise in the following way [Penrose (1997)]. If only slit A were open and the other closed, there would be many places on the screen which the photon could reach. If we now close the slit A and open the slit B, we may again find that the photon could reach the same spot on the screen.
However, if we open both slits, and if we have chosen the point on the screen carefully, we may now find that the photon cannot reach that spot, even though it could have done so if either slit alone were open. Somehow, the two possible things which the photon might do cancel each other out. This type of behavior does not take place in classical physics. Either one thing happens or another thing happens – we do not get two possible things which might happen, somehow conspiring to cancel each other out.

Fig. 1.1 The two–slit experiment, with individual photons of monochromatic light (see text for explanation).

The way we understand the outcome of this experiment in quantum mechanics is to say that when the photon is en route from the source to the screen, the state of the photon is not that of having gone through one slit or the other, but is some mysterious combination of the two, weighted by complex numbers [Penrose (1997)]. That is, we can write the state of the photon as a wave function \( \psi \), which is the linear superposition of the two states, \( |A\rangle \) and \( |B\rangle \), corresponding to the A–slot and B–slot alternatives,

\[
|\psi\rangle = z_1|A\rangle + z_2|B\rangle,
\]

where \( z_1 \) and \( z_2 \) are complex numbers (not both zero), while \( |\cdot\rangle \) denotes the quantum state ket–vector.

Now, in quantum mechanics, we are not so interested in the sizes of the

---

1In the Schrödinger picture, the unitary evolution \( U \) of a quantum system is described by the Schrödinger equation, which provides the time rate of change of the quantum state or wave function \( \psi = \psi(t) \).

2We are using here the standard Dirac ‘bra–ket’ notation for quantum states. Paul Dirac was one of the outstanding physicists of the 20th century. Among his achievements was a general formulation of quantum mechanics (having Heisenberg matrix mechanics and Schrödinger wave mechanics as special cases) and also its relativistic generalization involving the ‘Dirac equation’, which he discovered, for the electron. He had an unusual ability to ‘smell out’ the truth, judging his equations, to a large degree, by their aesthetic qualities!
complex numbers \( z_1 \) and \( z_2 \) themselves as we are in their ratio – it is only
the ratio of these numbers which has direct physical meaning (as multiplying
a quantum state with a nonzero complex number does not change the
physical situation). Recall that the Riemann sphere is a way of representing
complex numbers (plus \( \infty \)) and their ratios on a sphere on unit radius,
whose equatorial plane is the complex–plane, whose center is the origin of
that plane and the equator of this sphere is the unit circle in the complex–
plane. We can project each point on the equatorial complex–plane onto the
Riemann sphere, projecting from its south pole \( S \), which corresponds to the
point at infinity in the complex–plane. To represent a particular complex
ratio, say \( u = z/w \) (with \( w \neq 0 \)), we take the stereographic projection from
the sphere onto the plane.

The Riemann sphere plays a fundamental role in the quantum picture
of two–state systems [Penrose (1994)]. If we have a spin–\( \frac{1}{2} \) particle, such as
an electron, a proton, or a neutron, then the various combinations of their
spin states can be realised geometrically on the Riemann sphere. Spin –\( \frac{1}{2} \)
particles can have two spin states: (i) spin–up (with the rotation vector pointing upwards), and (ii) spin–down (with the rotation vector pointing downwards). The superposition of the two spin–states can be represented
symbolically as

\[
\lvert \uparrow \rangle = w \lvert \uparrow \rangle + z \lvert \downarrow \rangle.
\]

Different combinations of these spin states give us rotation about some
other axis and, if we want to know where that axis is, we take the ratio of
complex numbers \( u = z/w \). We place this new complex number \( u \) on the
Riemann sphere and the direction of \( u \) from the center is the direction of the
spin axis (see Figure 1.2).

More general quantum state vectors might have a form such as [Penrose
(1994)]:

\[
\lvert \psi \rangle = z_1 \lvert A_1 \rangle + z_2 \lvert A_2 \rangle + \ldots + z_n \lvert A_n \rangle,
\]

where \( z_1 \ldots z_n \) are complex numbers (not all zero) and the state vectors
\( \lvert A_1 \rangle, \ldots, \lvert A_n \rangle \) might represent various possible locations for a particle (or
perhaps some other property of a particle, such as its state of spin). Even
more generally, infinite sums would be allowed for a wave \( \psi \)–function or
quantum state vector.

Now, the most basic feature of unitary quantum evolution \( U \) is that it

\footnote{Recall that unitary quantum evolution \( U \) is governed by the time–dependent}
is linear. This means that, if we have two states, say $|\psi\rangle$ and $|\phi\rangle$, and if the Schrödinger equation would tell us that, after some time $t$, the states $|\psi\rangle$ and $|\phi\rangle$ would each individually evolve to new states $|\psi'\rangle$ and $|\phi'\rangle$, respectively then any linear superposition $z_1|\psi\rangle + z_2|\phi\rangle$, must evolve, after some time $t$, to the corresponding superposition $z_1|\psi'\rangle + z_2|\phi'\rangle$. Let us use the symbol $\mapsto$ to denote the evolution after time $t$, Then linearity asserts that if

$$|\psi\rangle \mapsto |\psi'\rangle \quad \text{and} \quad |\phi\rangle \mapsto |\phi'\rangle,$$

then the evolution

$$z_1|\psi\rangle + z_2|\phi\rangle \mapsto z_1|\psi'\rangle + z_2|\phi'\rangle$$

would also hold. This would consequently apply also to linear superpositions of more than two individual quantum states. For example, the

\textit{Schrödinger equation},

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H|\psi(t)\rangle,$$

where $\hbar \equiv \partial/\partial t$, $\hbar$ is the Planck’s constant, and $H$ is the Hamiltonian (total energy) operator. Given the quantum state $|\psi(t)\rangle$ at some initial time ($t = 0$), we can integrate the Schrödinger equation to get the state at any subsequent time. In particular, if $H$ is independent of time, then

$$|\psi(t)\rangle = \exp \left( -\frac{iHt}{\hbar} \right) |\psi(0)\rangle.$$
$z_1|\psi\rangle + z_2|\phi\rangle + z_3|\chi\rangle$ would evolve, after time $t$, to $z_1|\psi'\rangle + z_2|\phi'\rangle + z_3|\chi'\rangle$, if $|\psi\rangle$, $|\phi\rangle$, and $|\chi\rangle$ would each individually evolve to $|\psi'\rangle$, $|\phi'\rangle$, and $|\chi'\rangle$, respectively. Thus, the evolution always proceeds as though each different component of a superposition were oblivious to the presence of the others.

As a second experiment, consider a situation in which light impinges on a half–silvered mirror, that is a semi-transparent mirror that reflects just half the light (composed of a stream of photons) falling upon it and transmits the remaining half [Penrose (1994)]. We might well have imagined that for a stream of photons impinging on our half–silvered mirror, half the photons would be reflected and half would be transmitted. Not so! Quantum theory tells us that, instead, each individual photon, as it impinges on the minor, is separately put into a superposed state of reflection and transmission. If the photon before its encounter with the minor is in state $|A\rangle$, then afterwards it evolves according to $U$ to become a state that can be written $|B\rangle + i|C\rangle$, where $|B\rangle$ represents the state in which the photon is transmitted through the mirror and $|C\rangle$ the state where the photon is reflected from it (see Figure 1.3). Let us write this as

$|A\rangle \rightarrow |B\rangle + i|C\rangle$.

The imaginary factor ‘i’ arises here because of a net phase shift by a quarter of a wavelength (see [Klein and Furtak (1986)]), which occurs between the reflected and transmitted beams at such a mirror.

Although, from the classical picture of a particle, we would have to imagine that $|B\rangle$ and $|C\rangle$ just represent alternative things that the photon might do, in quantum mechanics we have to try to believe that the photon is now actually doing both things at once in this strange, complex superposition. To see that it cannot just be a matter of classical probability–weighted alternatives, let us take this example a little further and try to bring the
two parts of the photon state, i.e., the two photon beams, back together again [Penrose (1994)]. We can do this by first reflecting each beam with a fully silvered mirror. After reflection, the photon state $|B\rangle$ would evolve according to $U$, into another state $i|D\rangle$, whilst $|C\rangle$ would evolve into $i|E\rangle$,

$$
|B\rangle \rightarrow i|D\rangle \quad \text{and} \quad |C\rangle \rightarrow i|E\rangle.
$$

Thus the entire state $|B\rangle + i|C\rangle$ evolves by $U$ into

$$
|B\rangle + i|C\rangle \rightarrow i|D\rangle + i(i|E\rangle) = i|D\rangle - |E\rangle,
$$

since $i^2 = -1$. Now, suppose that these two beams come together at a fourth mirror, which is now half silvered (see Figure 1.4). The state $|D\rangle$ evolves into a combination $|G\rangle + i|F\rangle$, where $|G\rangle$ represents the transmitted state and $|F\rangle$ the reflected one. Similarly, $|E\rangle$ evolves into $|F\rangle + i|G\rangle$, since it is now the state $|F\rangle$ that is the transmitted state and $|G\rangle$ the reflected one,

$$
|D\rangle \rightarrow |G\rangle + i|F\rangle \quad \text{and} \quad |E\rangle \rightarrow |F\rangle + i|G\rangle.
$$

Our entire state $i|D\rangle - |E\rangle$ is now seen (because of the linearity of $U$) to evolve as:

$$
i|D\rangle - |E\rangle \rightarrow i(|G\rangle + i|F\rangle) - (|F\rangle + i|G\rangle)
$$

$$
= i|G\rangle - |F\rangle - |F\rangle - i|G\rangle = -2|F\rangle.
$$

As mentioned above, the multiplying factor $-2$ appearing here plays no physical role, thus we see that the possibility $|G\rangle$ is not open to the photon; the two beams together combine to produce just a single possibility $|F\rangle$. This curious outcome arises because both beams are present simultaneously in the physical state of the photon, between its encounters with the first and last mirrors. We say that the two beams interfere with one another.\footnote{This is a property of single photons: each individual photon must be considered to feel out both routes that are open to it, but it remains one photon; it does not split into two photons in the intermediate stage, but its location undergoes the strange kind of complex–number–weighted co–existence of alternatives that is characteristic of quantum theory.}
Fig. 1.4 *Mach–Zehnder interferometer*: the two parts of the photon state are brought together by two fully silvered mirrors (black), so as to encounter each other at a final half-silvered mirror (white). They interfere in such a way that the entire state emerges in state $|F\rangle$, and the detector at $G$ cannot receive the photon (see text for explanation).

### 1.2 Hilbert Space

#### 1.2.1 Quantum Hilbert Space

The family of all possible states ($|\psi\rangle$, $|\phi\rangle$, etc.) of a quantum system configure what is known as a *Hilbert space*. It is a complex vector space, which means that can perform the complex-number-weighted combinations that we considered before for quantum states. If $|\psi\rangle$ and $|\phi\rangle$ are both elements of the Hilbert space, then so also is $w|\psi\rangle + z|\phi\rangle$, for any pair of complex numbers $w$ and $z$. Here, we even allow $w = z = 0$, to give the element $0$ of the Hilbert space, which does not represent a possible physical state. We have the normal algebraic rules for a vector space:

\[
|\psi\rangle + |\phi\rangle = |\psi\rangle + |\phi\rangle, \\
|\psi\rangle + (|\phi\rangle + |\chi\rangle) = (|\psi\rangle + |\phi\rangle) + |\chi\rangle, \\
w(z|\psi\rangle) = (wz)|\psi\rangle, \\
(w + z)|\psi\rangle = w|\psi\rangle + z|\psi\rangle, \\
z(|\psi\rangle + |\phi\rangle) = z|\psi\rangle + z|\phi\rangle \\
0|\psi\rangle = 0, \quad z0 = 0.
\]

A Hilbert space can sometimes have a finite number of dimensions, as in the case of the spin states of a particle. For spin $\frac{1}{2}$, the Hilbert space is just 2D, its elements being the complex linear combinations of the two states $|\uparrow\rangle$ and $|\downarrow\rangle$. For spin $\frac{1}{2}n$, the Hilbert space is $(n+1)$D. However, sometimes the
Hilbert space can have an infinite number of dimensions, as e.g., the states of position or momentum of a particle. Here, each alternative position (or momentum) that the particle might have counts as providing a separate dimension for the Hilbert space. The general state describing the quantum location (or momentum) of the particle is a complex-number superposition of all these different individual positions (or momenta), which is the wave $\psi$–function for the particle.

Another property of the Hilbert space, crucial for quantum mechanics, is the Hermitian inner (scalar) product, which can be applied to any pair of Hilbert–space vectors to produce a single complex number. To understand how important the Hermitian inner product is for quantum mechanics, recall that the Dirac’s ‘bra–ket’ notation is formulated on its basis. If we have the two quantum states (i.e., Hilbert–space vectors) are $|\psi\rangle$ and $|\phi\rangle$, then their Hermitian scalar product is denoted $\langle\psi|\phi\rangle$, and it satisfies a number of simple algebraic properties:

\[
\langle\psi|\phi\rangle = \langle\phi|\psi\rangle, \quad \text{(bar denotes complex–conjugate)}
\]
\[
\langle\psi|(\phi + \chi)\rangle = \langle\psi|\phi\rangle + \langle\psi|\chi\rangle,
\]
\[
(z\langle\psi|)\phi = z\langle\psi|\phi\rangle,
\]
\[
\langle\psi|\phi\rangle \geq 0, \quad \langle\psi|\phi\rangle = 0 \quad \text{if} \quad |\psi\rangle = 0.
\]

For example, probability of finding a quantum particle at a given location is a squared length $|\psi|^2$ of a Hilbert–space position vector $|\psi\rangle$, which is the scalar product $\langle\psi|\psi\rangle$ of the vector $|\psi\rangle$ with itself. A normalized state is given by a Hilbert–space vector whose squared length is unity.

The second important thing that the Hermitian scalar product gives us is the notion of orthogonality between Hilbert–space vectors, which occurs when the scalar product of the two vectors is zero. In ordinary terms, orthogonal states are things that are independent of one another. The importance of this concept for quantum physics is that the different alternative outcomes of any measurement are always orthogonal to each other. For example, states $|\uparrow\rangle$ and $|\downarrow\rangle$ are mutually orthogonal. Also, orthogonal are all different possible positions that a quantum particle might be located in [Penrose (1994)].

### 1.2.2 Formal Hilbert Space

A norm on a complex vector space $\mathcal{H}$ is a mapping from $\mathcal{H}$ into the complex numbers, $\|\cdot\|: \mathcal{H} \to \mathbb{C}; h \mapsto \|h\|$, such that the following set of norm–axioms
hold:
(N1) \( \| h \| \geq 0 \) for all \( h \in \mathcal{H} \) and \( \| h \| = 0 \) implies \( h = 0 \) (positive definiteness);
(N2) \( \| \lambda h \| = |\lambda| \| h \| \) for all \( h \in \mathcal{H} \) and \( \lambda \in \mathbb{C} \) (homogeneity); and
(N3) \( \| h_1 + h_2 \| \leq \| h_1 \| + \| h_2 \| \) for all \( h_1, h_2 \in \mathcal{H} \) (triangle inequality).

The pair \( (\mathcal{H}, \| \cdot \|) \) is called a normed space.

A Hermitian inner product on a complex vector space \( \mathcal{H} \) is a mapping \( \langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C} \) such that the following set of inner–product–axioms hold:

(IP1) \( \langle h_1 + h_2, h \rangle = \langle h_1, h \rangle + \langle h_2, h \rangle \);
(IP2) \( \langle \alpha h, h_1 \rangle = \alpha \langle h, h_1 \rangle \);
(IP3) \( \langle h_1, h_2 \rangle = \overline{\langle h_2, h_1 \rangle} \) (so \( \langle h, h \rangle \) is real);
(IP4) \( \langle h, h \rangle \geq 0 \) and \( \langle h, h \rangle = 0 \) provided \( h = 0 \).

These properties are to hold for all \( h, h_1, h_2 \in \mathcal{H} \) and \( \alpha \in \mathbb{C} \); \( \bar{z} \) denotes the complex conjugate of the complex number \( z \). (IP2) and (IP3) imply that \( \langle \alpha h, h_1 \rangle = \bar{\alpha} \langle h, h_1 \rangle \). As is customary, for a complex number \( z \) we shall denote by \( \text{Re} z = \frac{z + \bar{z}}{2} \) and \( \text{Im} z = \frac{z - \bar{z}}{2i} \), \( |z| = (z\bar{z})^{1/2} \) its real and imaginary parts and its absolute value.

The standard inner product on the product space \( \mathbb{C}^n = \mathbb{C} \times \cdots \times \mathbb{C} \) is defined by \( \langle z, w \rangle = \sum_{i=1}^{n} z_i w_i \), and axioms (IP1)–(IP4) are readily checked. Also \( \mathbb{C}^n \) is a normed space with \( \| z \|^2 = \sum_{i=1}^{n} |z_i|^2 \).

The pair \( (\mathcal{H}, \langle \cdot, \cdot \rangle) \) is called an inner product space.

In an inner product space \( \mathcal{H} \), two vectors \( h_1, h_2 \in \mathcal{H} \) are called orthogonal, and we write \( h_1 \perp h_2 \), provided \( \langle h_1, h_2 \rangle = 0 \). For a subset \( A \subset \mathcal{H} \), the set \( A^\perp \) defined by \( A^\perp = \{ h \in \mathcal{H} | \langle h, x \rangle = 0 \text{ for all } x \in A \} \) is called the orthogonal complement of \( A \).

In an inner product space \( \mathcal{H} \) the Cauchy–Schwartz inequality holds:
\[
|\langle h_1, h_2 \rangle| \leq \langle h_1, h_1 \rangle^{1/2} \langle h_2, h_2 \rangle^{1/2}.
\]
Here, equality holds provided \( h_1, h_2 \) are linearly dependent.

Let \( (\mathcal{H}, \| \cdot \|) \) be an inner product space and set \( \| h \| = \langle h, h \rangle^{1/2} \). Then \( (\mathcal{H}, \| \cdot \|) \) is a normed space.

Let \( (\mathcal{H}, \langle \cdot, \cdot \rangle) \) be an inner product space and \( \| \cdot \| \) the corresponding norm. Then we have

1. **Polarization law:**
\[
4 \langle h_1, h_2 \rangle = \| h_1 + h_2 \|^2 - \| h_1 - h_2 \|^2 + i \| h_1 + ih_2 \|^2 - i \| h_1 - ih_2 \|^2,
\]
(2) **Parallelogram law:**

\[
2 \|h_1\|^2 + 2 \|h_2\|^2 = \|h_1 + h_2\|^2 - \|h_1 - h_2\|^2.
\]

Let \((\mathcal{H}, \|\cdot\|)\) be a normed space and define \(d(h_1, h_2) = \|h_1 - h_2\|\). Then \((\mathcal{H}, d)\) is a metric space.

Let \((\mathcal{H}, \|\cdot\|)\) be a normed space. If the corresponding metric \(d\) is complete, we say \((\mathcal{H}, \|\cdot\|)\) is a Banach space. If \((\mathcal{H}, \|\cdot\|)\) is an inner product space whose corresponding metric is complete, we say \((\mathcal{H}, \|\cdot\|)\) is a Hilbert space (see, e.g., [Abraham et al. (1988)]).

If \(\mathcal{H}\) is a Hilbert space and \(F\) it closed subspace, then \(\mathcal{H}\) splits into two mutually orthogonal subspaces, \(\mathcal{H} = F \oplus F^\perp\), where \(\oplus\) denotes the orthogonal sum. Thus every closed subspace of a Hilbert space splits.

Let \(\mathcal{H}\) be a Hilbert space. A set \(\{h_i\}_{i \in I}\) is called orthonormal if \(\langle h_i, h_j \rangle = \delta_{ij}\), the Kronecker delta. An orthonormal set \(\{h_i\}_{i \in I}\) is a Hilbert basis if \(\text{closure}(\text{span}\{h_i\}_{i \in I}) = \mathcal{H}\). Any Hilbert space has a Hilbert basis.

In the finite dimensional case equivalence and completeness are automatic. Let \(\mathcal{H}\) be a finite–dimensional vector space. Then (i) there is a norm on \(\mathcal{H}\); (ii) all norms on \(\mathcal{H}\) are equivalent; (iii) all norms on \(\mathcal{H}\) are complete.

Consider the space \(L^2([a,b], \mathbb{C})\) of square–Lebesgue–integrable complex–valued functions defined on an interval \([a,b] \subset \mathbb{C}\), that is, functions \(f\) that satisfy \(\int_a^b |f(x)|^2 \, dx < \infty\). It is a Banach space with the norm defined by \(\|f\| = \left( \int_a^b |f(x)|^2 \, dx \right)^{1/2}\), and a Hilbert space with the inner product defined by \(\langle f, g \rangle = \int_a^b f(x) \overline{g(x)} \, dx\).

Recall from elementary linear algebra that the dual space of a finite dimensional vector space of dimension \(n\) also has dimension \(n\) and so the space and its dual are isomorphic. It is also true for Hilbert space.

**Riesz Representation Theorem.** Let \(\mathcal{H}\) be be a real (resp., complex) Hilbert space. The map \(h \mapsto \langle \cdot, h \rangle\) is a linear (resp., antilinear) norm–preserving isomorphism of \(\mathcal{H}\) with \(\mathcal{H}^*\); for short, \(\mathcal{H} \cong \mathcal{H}^*\). (A map \(A : \mathcal{H} \to F\) between complex vector spaces is called antilinear if we have the identities \(A(h + h') = Ah + Ah'\), and \(A(\alpha h) = \bar{\alpha}Ah\).)

Let \(\mathcal{H}\) and \(F\) be Banach spaces. We say \(\mathcal{H}\) and \(F\) are in strong duality if there is a non–degenerate continuous bilinear functional \(\langle \cdot, \cdot \rangle : \mathcal{H} \times F \to \mathbb{R}\), also called a pairing of \(\mathcal{H}\) with \(F\). Now, let \(\mathcal{H} = F\) and \(\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{R}\) be an inner product on \(\mathcal{H}\). If \(\mathcal{H}\) is a Hilbert space, then \(\langle \cdot, \cdot \rangle\) is a strongly non–degenerate pairing by the Riesz representation Theorem.
1.3 Human Intelligence, Mind and Reason

Recall that the word intelligence (plural intelligences) comes from Latin intelligentia.\footnote{Intelligentia is a combination of Latin inter = between and legere = choose, pick out, read. Inter–lege–nt–ia, literally means ‘choosing between.’ Also, note that there is a is a scientific journal titled ‘Intelligence’, dealing with intelligence and psychometrics. It was founded in 1977 by Douglas K. Detterman of Case Western Reserve University. It is currently published by Elsevier and is the official journal of the International Society for Intelligence Research.} It is a property of human mind that encompasses many related mental abilities, such as the capacities to reason, plan, solve problems, think abstractly, comprehend ideas and language, and learn. Although many regard the concept of intelligence as having a much broader scope, for example in cognitive science and computer science, in some schools of psychology\footnote{Recall that psychology is an academic and applied field involving the study of the human mind, brain, and behavior. Psychology also refers to the application of such knowledge to various spheres of human activity, including problems of individuals’ daily lives and the treatment of mental illness. Psychology differs from anthropology, economics, political science, and sociology in seeking to explain the mental processes and behavior of individuals. Psychology differs from biology and neuroscience in that it is primarily concerned with the interaction of mental processes and behavior, and of the overall processes of a system, and not simply the biological or neural processes themselves, though the subfield of neuropsychology combines the study of the actual neural processes with the study of the mental effects they have subjectively produced. The word psychology comes from the ancient Greek ‘psyche’, which means ‘soul’ or ‘mind’ and ‘ology’, which means ‘study’.

Also, note that there is a is a scientific journal titled ‘Intelligence’, dealing with intelligence and psychometrics. It was founded in 1977 by Douglas K. Detterman of Case Western Reserve University. It is currently published by Elsevier and is the official journal of the International Society for Intelligence Research.}, the study of intelligence generally regards this trait as distinct from creativity, personality, character, or wisdom.

Briefly, the word intelligence has five common meanings:

(1) Capacity of human mind, especially to understand principles, truths, concepts, facts or meanings, acquire knowledge, and apply it to practise; the ability to learn and comprehend.

(2) A form of life that has such capacities.

(3) Information, usually secret, about the enemy or about hostile activities.

(4) A political or military department, agency or unit designed to gather such information.

(5) Biological intelligent behavior represents animal’s ability to make productive decisions for a specific task, given a root objective; this decision is based on learning which requires the ability to hold onto results from previous tasks, as well as being able to analyze the situation; the root objective for living organisms is simply survival; the ‘specific task’ could
be a choice of food, i.e., one that provides long steady supply of energy as it could be a long while before the next mealtime; this is in perfect harmony with the root biological objective – survival.

According to Encyclopedia Britannica, intelligence is the ability to adapt effectively to the environment, either by making a change in oneself or by changing the environment or finding a new one. Different investigators have emphasized different aspects of intelligence in their definitions. For example, in a 1921 symposium on the definition of intelligence, the American psychologist Lewis Terman emphasized the ability to think abstractly, while another American psychologist, Edward Thorndike, emphasized learning and the ability to give good responses to questions. In a similar 1986 symposium, however, psychologists generally agreed on the importance of adaptation to the environment as the key to understanding both what intelligence is and what it does. Such adaptation may occur in a variety of environmental situations. For example, a student in school learns the material that is required to pass or do well in a course; a physician treating a patient with an unfamiliar disease adapts by learning about the disease; an artist reworks a painting in order to make it convey a more harmonious impression. For the most part, adapting involves making a change in oneself in order to cope more effectively, but sometimes effective adaptation involves either changing the environment or finding a new environment altogether. Effective adaptation draws upon a number of cognitive processes, such as perception, learning, memory, reasoning, and problem solving. The main trend in defining intelligence, then, is that it is not itself a cognitive or mental process, but rather a selective combination of these processes purposively directed toward effective adaptation to the environment. For example, the physician noted above learning about a new disease adapts by perceiving material on the disease in medical literature, learning what the material contains, remembering crucial aspects of it that are needed to treat the patient, and then reasoning to solve the problem of how to apply the information to the needs of the patient. Intelligence, in sum, has come to be regarded as not a single ability but an effective drawing together of many abilities. This has not always been obvious to investigators of the subject, however, and, indeed, much of the history of the field revolves around arguments regarding the nature and abilities that constitute intelligence.
Human Reason

Recall that in the philosophy of arguments, reason is the ability of the human mind to form and operate on concepts in abstraction, in varied accordance with rationality and logic — terms with which reason shares heritage. Reason is thus a very important word in Western intellectual history, to describe a type or aspect of mental thought which has traditionally been claimed as distinctly human, and not to be found elsewhere in the animal world. Discussion and debate about the nature, limits and causes of reason could almost be said to define the main lines of historical philosophical discussion and debate. Discussion about reason especially concerns:

(a) its relationship to several other related concepts: language, logic, consciousness etc,
(b) its ability to help people decide what is true, and
(c) its origin.

The concept of reason is connected to the concept of language, as reflected in the meanings of the Greek word ‘logos’, later to be translated by Latin ‘ratio’ and then French ‘raison’, from which the English word derived. As reason, rationality, and logic are all associated with the ability of the human mind to predict effects as based upon presumed causes, the word ‘reason’ also denotes a ground or basis for a particular argument, and hence is used synonymously with the word ‘cause’.

It is sometimes said that the contrast between reason and logic extends back to the time of Plato and Aristotle. Indeed, although they had no

---

5 Plato (c. 427 — c. 347 BC) was an immensely influential ancient Greek philosopher, a student of Socrates, writer of philosophical dialogues, and founder of the Academy in Athens where Aristotle studied. Plato lectured extensively at the Academy, and wrote on many philosophical issues, dealing especially in politics, ethics, metaphysics, and epistemology. The most important writings of Plato are his dialogues, although some letters have come down to us under his name. It is believed that all of Plato’s authentic dialogues survive. However, some dialogues ascribed to Plato by the Greeks are now considered by the consensus of scholars to be either suspect (e.g., First Alcibiades, Clitophon) or probably spurious (such as Demodocus, or the Second Alcibiades). The letters are all considered to probably be spurious, with the possible exception of the Seventh Letter. Socrates is often a character in Plato’s dialogues. How much of the content and argument of any given dialogue is Socrates’ point of view, and how much of it is Plato’s, is heavily disputed, since Socrates himself did not write anything; this is often referred to as the ‘Socratic problem’. However, Plato was doubtless strongly influenced by Socrates’ teachings.

Platonism has traditionally been interpreted as a form of metaphysical dualism, sometimes referred to as Platonic realism, and is regarded as one of the earlier representatives of metaphysical objective idealism. According to this reading, Plato’s metaphysics
separate Greek word for logic as opposed to language and reason, Aristotle’s *syllogism* (Greek ‘syllogismos’) identified logic clearly for the first time as a distinct field of study: the most peculiarly reasonable (‘logikê’) part of reasoning, so to speak.

No philosopher of any note has ever argued that logic is the same as reason. They are generally thought to be distinct, although logic is one important aspect of reason. But the tendency to the preference for ‘hard logic’, or ‘solid logic’, in modern times has incorrectly led to the two terms occasionally being seen as essentially synonymous or perhaps more often logic is seen as the defining and pure form of reason.

However, machines and animals can unconsciously perform logical operations, and many animals (including humans) can unconsciously, associate different perceptions as causes and effects and then make decisions or even plans. Therefore, to have any distinct meaning at all, ‘reason’ must be the type of thinking which links language, consciousness and logic, and at this time, only humans are known to combine these things.

However, note that reasoning is defined very differently depending on the context of the understanding of reason as a form of knowledge. The logical definition is the act of using reason to derive a conclusion from certain premises using a given methodology, and the two most commonly used explicit methods to reach a conclusion are deductive reasoning and inductive reasoning. However, within idealist philosophical contexts, reasoning is the mental process which informs our imagination, perceptions, thoughts, and feelings with whatever intelligibility these appear to contain; and thus links our experience with universal meaning. The specifics of the methods of reasoning are of interest to such disciplines as philosophy, logic, psychology, and artificial intelligence.

In deductive reasoning, given true premises, the conclusion must follow and it cannot be false. In this type of reasoning, the conclusion is inherent in the premises. Deductive reasoning therefore does not increase one’s knowledge base and is said to be non–ampliative. Classic examples of deductive reasoning are found in such syllogisms as the following:

\[
\text{divides the world into two distinct aspects: the } \textit{intelligible world} \text{ of ‘forms’, and the } \textit{perceptual world} \text{ we see around us. The perceptual world consists of imperfect copies of the intelligible forms or ideas. These forms are unchangeable and perfect, and are only comprehensible by the use of the intellect or understanding, that is, a capacity of the mind that does not include sense-perception or imagination. This division can also be found in Zoroastrian philosophy, in which the dichotomy is referenced as the } \textit{Minu} (intelligence) \text{ and } \textit{Giti} (perceptual) \text{ worlds. Currently, in the domain of mathematical physics, this view has been adopted by Sir Roger Penrose} \text{ [Penrose (1967)].} \\
\]
(1) One must exist/live to perform the act of thinking.
(2) I think.
(3) Therefore, I am.

Aristotle (384 BC — March 7, 322 BC) was an ancient Greek philosopher, a student of Plato and teacher of Alexander the Great. He wrote books on diverse subjects, including physics, poetry, zoology, logic, rhetoric, government, and biology, none of which survive in their entirety. Aristotle, along with Plato and Socrates, is generally considered one of the most influential of ancient Greek philosophers. They transformed Presocratic Greek philosophy into the foundations of Western philosophy as we know it. The writings of Plato and Aristotle founded two of the most important schools of Ancient philosophy.

Aristotle valued knowledge gained from the senses and in modern terms would be classed among the modern empiricists. He also achieved a ‘grounding’ of dialectic in the Topics by allowing interlocutors to begin from commonly held beliefs (Endoxa), with his frequent aim being to progress from ‘what is known to us’ towards ‘what is known in itself’ (Physics). He set the stage for what would eventually develop into the empirical scientific method some two millennia later. Although he wrote dialogues early in his career, no more than fragments of these have survived. The works of Aristotle that still exist today are in treatise form and were, for the most part, unpublished texts. These were probably lecture notes or texts used by his students, and were almost certainly revised repeatedly over the course of years. As a result, these works tend to be eclectic, dense and difficult to read. Among the most important ones are Physics, Metaphysics (or Ontology), Nicomachean Ethics, Politics, De Anima (On the Soul) and Poetics. These works, although connected in many fundamental ways, are very different in both style and substance.

Aristotle is known for being one of the few figures in history who studied almost every subject possible at the time, probably being one of the first polymaths. In science, Aristotle studied anatomy, astronomy, economics, embryology, geography, geology, meteorology, physics, and zoology. In philosophy, Aristotle wrote on aesthetics, ethics, government, metaphysics, politics, psychology, rhetoric and theology. He also dealt with education, foreign customs, literature and poetry. His combined works practically constitute an encyclopedia of Greek knowledge.

According to Aristotle, everything is made out of the five basic elements:

(1) Earth, which is cold and dry;
(2) Water, which is cold and wet;
(3) Fire, which is hot and dry;
(4) Air, which is hot and wet; and
(5) Aether, which is the divine substance that makes up the heavenly spheres and heavenly bodies (stars and planets).

Aristotle defines his philosophy in terms of essence, saying that philosophy is ‘the science of the universal essence of that which is actual’. Plato had defined it as the ‘science of the idea’, meaning by idea what we should call the unconditional basis of phenomena. Both pupil and master regard philosophy as concerned with the universal; Aristotle, however, finds the universal in particular things, and called it the essence of things, while Plato finds that the universal exists apart from particular things, and is related to them as their prototype or exemplar. For Aristotle, therefore, philosophic
In inductive reasoning, on the other hand, when the premises are true, then the conclusion follows with some degree of probability. This method implies the ascent from the study of particular phenomena to the knowledge of essences, while for Plato philosophic method means the descent from a knowledge of universal ideas to a contemplation of particular imitations of those ideas. In a certain sense, Aristotle’s method is both inductive and deductive, while Plato’s is essentially deductive from a priori principles.

In the larger sense of the word, Aristotle makes philosophy coextensive with reasoning, which he also called ‘science’. Note, however, that his use of the term science carries a different meaning than that which is covered by the scientific method. “All science (dianoia) is either practical, poetical or theoretical.” By practical science he understands ethics and politics; by poetical, he means the study of poetry and the other fine arts; while by theoretical philosophy he means physics, mathematics, and metaphysics.

Aristotle’s conception of logic was the dominant form of logic up until the advances in mathematical logic in the 19th century. Kant himself thought that Aristotle had done everything possible in terms of logic. The Organon is the name given by Aristotle’s followers, the Peripatetics, for the standard collection of six of his works on logic. The system of logic described in two of these works, namely On Interpretation and the Prior Analytics, is often called Aristotelian logic.

Aristotle was the creator of syllogisms with modalities (modal logic). The word modal refers to the word ‘modes’, explaining the fact that modal logic deals with the modes of truth. Aristotle introduced the qualification of ‘necessary’ and ‘possible’ premises. He constructed a logic which helped in the evaluation of truth but which was difficult to interpret.

Recall that the word probability derives from the Latin ‘probare’ (to prove, or to test). Informally, probable is one of several words applied to uncertain events or knowledge, being closely related in meaning to likely, risky, hazardous, and doubtful. Chance, odds, and bet are other words expressing similar notions. Just as the theory of mechanics assigns precise definitions to such everyday terms as work and force, the theory of probability attempts to quantify the notion of probable.

The scientific study of probability is a modern development. Gambling shows that there has been an interest in quantifying the ideas of probability for millennia, but exact mathematical descriptions of use in those problems only arose much later. The doctrine of probabilities dates to the correspondence of Pierre de Fermat and Blaise Pascal (1654). Christiaan Huygens (1657) gave the earliest known scientific treatment of the subject. Jakob Bernoulli’s ‘Ars Conjectandi’ (posthumous, 1713) and Abraham de Moivre’s ‘Doctrine of Chances’ (1718) treated the subject as a branch of mathematics.

Pierre-Simon Laplace (1774) made the first attempt to deduce a rule for the combination of observations from the principles of the theory of probabilities. He represented the law of probability of errors by a curve $y = \varphi(x)$, $x$ being any error and $y$ its probability, and laid down three properties of this curve: (i) it is symmetric as to the $y$–axis; (ii) the $x$–axis is an asymptote, the probability of the error being 0; (iii) the area enclosed is 1, it being certain that an error exists. He deduced a formula for the mean of three observations. He also gave (1781) a formula for the law of facility of error (a term due to Lagrange, 1774), but one which led to unmanageable equations. Daniel Bernoulli (1778) introduced the principle of the maximum product of the probabilities of a system of concurrent errors.

The method of least squares is due to Adrien–Marie Legendre (1805), who introduced it in his ‘Nouvelles méthodes pour la détermination des orbites des comètes’ (New Methods
of reasoning is ampliative, as it gives more information than what was contained in the premises themselves. A classical example comes from David Hume:  

(1) The sun rose in the east every morning up until now.
(2) Therefore the sun will also rise in the east tomorrow.

A third method of reasoning is called abductive reasoning, or inference to the best explanation. This method is more complex in its structure and can involve both inductive and deductive arguments. The main characteristic of abduction is that it is an attempt to favor one conclusion above others by either attempting to falsify alternative explanations, or showing the likelihood of the favored conclusion given a set of more or less disputable assumptions.

A fourth method of reasoning is analogy. Reasoning by analogy goes from a particular to another particular. The conclusion of an analogy is only plausible. Analogical reasoning is very frequent in common sense, science, philosophy and the humanities, but sometimes it is accepted only as an auxiliary method. A refined approach is case–based reasoning.
This page intentionally left blank
Chapter 2

Elements of Quantum Mechanics

In this chapter we present the essence of quantum dynamics in a complex Hilbert space, mainly using the quantum formalism of P.A.M. Dirac.

2.1 Basics of Non–Relativistic Quantum Mechanics

Recall that Heisenberg, with his discovery of quantum mechanics (1925; see Cassidy (1992)), introduced a new outlook on the nature of physical theory. Previously, it was always considered essential that there should be a detailed description of what is taking place in natural phenomena, and one used this description to calculate results comparable with experiment. Heisenberg put forward the view that it is sufficient to have a mathematical scheme from which one can calculate in a consistent manner the results of all experiments. That is, a detailed description in the traditional sense is unnecessary and may very well be impossible to establish [Dirac (1928a)] [Dirac (1928b)] [Dirac (1926c)].

Heisenberg’s method focuses attention on the quantities which enter into experimental results. It was first applied to the spectral theory, for which these quantities are the energy levels of the atomic system and certain probability coefficients, which determine the probability of a radiative transition taking place from one level to another. The method sets up equations connecting these quantities and allows one to calculate them, but does not go beyond this. It does not provide any description of radiative transition processes. It does not even allow one to deduce how the results of a calculation are to be used, but requires one to assume Einstein’s laws of radiation (the laws which tell how the probability of a radiative transition process depends on the intensity of the incident radiation), and to assume that certain quantities determined by the calculation are the coefficients
appearing in the laws.

Shortly after Heisenberg’s discovery, Schrödinger set up independently another form of quantum mechanics (1926; see Moore (1989)), which also enables one to calculate energy levels and probability coefficients and gives results agreeing with those of Heisenberg, but which introduces an important new feature. It connects together, in one calculation, a set of probability coefficients that act together under certain conditions in Nature; e.g., the set of probability coefficients referring to transitions from one particular initial state to any final state. In this respect, Schrödinger’s method is to be contrasted with Heisenberg’s method, which connects together in one calculation all the probability coefficients for a dynamical system, i.e., the probability coefficients from all initial states to all final states.

This feature of Schrödinger’s method gives it two important advantages [Dirac (1925); Dirac (1926e)]. First, as a consequence of its enabling one to get fewer results at a time, it makes the computation much simpler. Secondly, it supplies, in a certain sense, a description of what is taking place in Nature, since a calculation leading to results that come into play together under certain conditions in Nature will be in close correspondence with the physical process that is taking place under those conditions, various points in the calculation having their counterparts in the physical process. A description in this limited sense seems to be the most that is possible for atomic processes. It implies a much less complete connection between the mathematics and the physics than one has in classical mechanics, and one might be disinclined to call it a description at all, but one may at least consider it as an appropriate generalization of what one usually means by a description. On account of Schrödinger’s method allowing a description in this new sense while Heisenberg’s allows none, Schrödinger’s method introduces an outlook on the nature of physical theory intermediate between Heisenberg’s and the old classical (Newton–Maxwellian) one.

When Heisenberg’s and Schrödinger’s theories were developed it was soon found by Dirac that they both rested on the same mathematical formalism and differed only with regard to the method of physical interpretation (see Dirac (1949)). Dirac’s formalism is a generalization of the Hamiltonian form of classical Newtonian dynamics, involving linear operators instead of ordinary algebraic variables, and is so natural and beautiful as to make one feel sure of its correctness as the foundation of the theory. The question of its interpretation, however, which involved unifying Heisenberg’s and Schrödinger’s ideas into a satisfactory comprehensive scheme, was not so easily settled.
The situation of a formalism (in this case, Dirac’s) becoming established before one is clear about its interpretation should not be considered as surprising, but rather as a natural consequence of the drastic alterations which the development of physics had required in some of the basic physical concepts. This made it an easier matter to discover the mathematical formalism needed for a fundamental physical theory than its interpretation, since the number of things one had to choose between in discovering the formalism was very limited, the number of fundamental ideas in pure mathematics being not very great, while with the interpretation most unexpected things might turn up.

The best way of seeking the interpretation in such cases is probably from a discussion of simple examples. This way was used for the theory of quantum mechanics and led eventually to a satisfactory interpretation applicable to all phenomena for which relativistic effects are negligible. This interpretation is more closely connected with Schrödinger’s method than Heisenberg’s, as one would expect on account of the former affording in some sense a description of Nature, and is centered round a Schrödinger’s wave $\psi$–function, which is one of the things that can be operated on by the linear operators which the dynamical variables have become. The correspondence which the existence of a description implies between the mathematics and the physics makes a wave $\psi$–function correspond to a state of motion of the atomic system, in such a way that, for example, a calculation which gives the transition probabilities from a particular initial state to any final state would be based on that wave $\psi$–function which represents the motion ensuing from this initial state. A wave $\psi$–function is a complex function $\psi(q_1, q_2, \ldots, q_n, t)$ of all the coordinates $q_1, q_2, \ldots, q_n, t$ of the system and of the time $t$, and it receives the interpretation that the square of its modulus, $|\psi(q_1, q_2, \ldots, q_n, t)|^2$, is the probability, for the state of motion it corresponds to, of the coordinates having values in the neighborhood of $q_1, q_2, \ldots, q_n$, per unit volume of coordinate space (or, configuration space), at the time $t$.

A wave $\psi$–function can be transformed so as to refer to other dynamical variables, for example, the momenta $p_1, p_2, \ldots, p_n$, when it is said to be in another representation. The square of its modulus $|\psi(p_1, p_2, \ldots, p_n, t)|^2$ is then the probability, per unit volume of momentum space (or, phase–space), of the momenta having values in the neighborhood of $p_1, p_2, \ldots, p_n$ at the time $t$. A wave $\psi$–function itself never has an interpretation, but only the square of its modulus, and the need for distinguishing between two wave functions having the same squares of their moduli arises only because,
if they are transformed to a different representation, the squares of their moduli will in general become different. This brings out the *incompleteness of description*, which is possible with quantum mechanics [Dirac (1928a); Dirac (1928b); Dirac (1926c); Dirac (1949)].

One may make a slight modification in the wave functions in any representation by introducing a weight factor \( \lambda \) and arranging for the probability to be \( \lambda |\psi|^2 \) instead of \( |\psi|^2 \). The weight factor may be any positive function of the variables occurring in the wave function.

Wave functions have to satisfy a certain *wave equation*, namely, the equation

\[
i\hbar \partial_t \psi = H \psi, \tag{2.1}\]

where \( \partial_t \equiv \partial/\partial t \), \( i = \sqrt{-1} \), \( \hbar \) is the Planck’s constant, and \( H \) is a *Hermitian (self–adjoint) linear operator* representing the Hamiltonian of the system (expressed in the representation concerned). The wave equation (2.1) is a generalization of the *Hamilton–Jacobi equation* of classical mechanics. If \( S \) is a solution of the latter equation, then

\[
\psi = e^{iS/\hbar} \tag{2.2}\]

will give a first approximation to a solution of the former.

An important property of the wave equation (2.1) is that it yields the *probability conservation law*: the total probability of the variables occurring in the wave \( \psi \)–function having any value is constant. The wave \( \psi \)–function should be normalized so as to make this probability initially unity and then it always remains unity. This conservation law is a mathematical consequence of the wave equation being linear in the operator \( \partial_t \) and of \( H \) being a self–adjoint operator.

The wave equation is linear and homogeneous in the wave \( \psi \)–function and so are the transformation equations. In consequence, one can add together two \( \psi \)'s and get a third. The correspondence between \( \psi \)'s and states of motion now allows one to infer that there is a relationship between the states of motion, such that one can add or superpose two states to get a third. This relationship constitutes the *Principle of superposition of states*, one of the general principles governing the interpretation of quantum mechanics.

Another of these principles is *Heisenberg uncertainty principle*. This is a consequence of the transformation laws connecting \( \psi(q) \) and \( \psi(p) \), which show that each of these functions is the *Fourier transform* of the other,
apart from numerical coefficients, so that one meets the same limitations in giving values to a $q$ and $p$ as in giving values to the position and frequency of a train of waves [Dirac (1926e) Dirac (1949)]. These general principles serve to bring out the departures needed from ordinary classical (Newton–Maxwellian) ideas. They are of so drastic and unexpected a nature that it is not to be wondered at that they were discovered only indirectly, as consequences of a previously established mathematical scheme, instead of being built up directly from experimental facts.

2.1.1 Canonical Quantization

To make a leap into the quantum realm, recall that classical state–space for the biodynamic system of $n$ point–particles is its 6ND phase–space $\mathcal{P}$, including all position and momentum vectors, $\mathbf{r}_i = (x, y, z)_i$ and $\mathbf{p}_i = (p_x, p_y, p_z)_i$, respectively, for $i = 1, ..., n$.

The quantization is performed as a linear representation of the real Lie algebra $\mathcal{L}_P$ of the phase–space $\mathcal{P}$, defined by the Poisson bracket $\{f, g\}$ of classical variables $f, g$ – into the corresponding real Lie algebra $\mathcal{L}_H$ of the Hilbert space $\mathcal{H}$, defined by the commutator $[\hat{f}, \hat{g}]$ of skew–Hermitian operators $\hat{f}, \hat{g}$. This sounds like a functor, however it is not; as J. Baez says, ‘First quantization is a mystery, but second quantization is a functor’. Mathematically, if quantization were natural it would be a functor from the category Symplec, whose objects are symplectic manifolds (i.e., phase–spaces) and whose morphisms are symplectic maps (i.e., canonical transformations) to the category Hilbert, whose objects are Hilbert spaces and whose morphisms are unitary operators.

Historically first, the so–called canonical quantization is based on the so–called Dirac rules for quantization. It is applied to ‘simple’ systems: finite number of degrees–of–freedom and ‘flat’ classical phase–spaces (an open set of $\mathbb{R}^{2n}$). Canonical quantization includes the following data [Dirac (1949)]:

1. **Classical description.** The system is described by the Hamiltonian or canonical formalism: its classical phase–space is locally coordinated by a set of canonical coordinates $(q^j, p_j)$, the position and momentum coordinates. Classical observables are real functions $f(q^j, p_j)$. Eventually, a Lie group $G$ of symmetries acts on the system.

2. **Quantum description.** The quantum phase–space is a complex Hilbert space $\mathcal{H}$. Quantum observables are Hermitian (i.e., self–adjoint) oper-
atners acting on $\mathcal{H}$. (The Hilbert space is complex in order to take into account the interference phenomena of wave functions representing the quantum states. The operators are self–adjoint in order to assure their eigenvalues are real.) The symmetries of the system are realized by a group of unitary operators $U_G(\mathcal{H})$.

(3) **Quantization method.** As a Hilbert space we take the space of square integrable complex functions of the configuration space; that is, functions depending only on the position coordinates, $\psi(q^i)$. The quantum operator associated with $f(q^i, p_j)$ is obtained by replacing $p_j$ by $-i\hbar \partial_{q^j}$, and hence we have the correspondence $f(q^i, p_j) \leftrightarrow \hat{f}(q^i, -i\hbar \partial_{q^j})$. In this way, the classical commutation rules between the canonical coordinates are assured to have a quantum counterpart: the commutation rules between the quantum operators of position and momentum (which are related to the ‘uncertainty principle’ of quantum mechanics).

### 2.1.2 Quantum States and Operators

Quantum systems have two modes of evolution in time. The first, governed by standard, **time–dependent Schrödinger equation**:

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = \hat{H} |\psi\rangle,$$  \hspace{1cm} \text{(2.3)}

describes the time evolution of quantum systems when they are undisturbed by measurements. ‘Measurements’ are defined as interactions of the quantum system with its classical environment. As long as the system is sufficiently isolated from the environment, it follows Schrödinger equation. If an interaction with the environment takes place, i.e., a measurement is performed, the system abruptly *decoheres* i.e., collapses or reduces to one of its classically allowed states.

A **time–dependent state of a quantum system** is determined by a normalized, complex, wave **psi–function** $\psi = \psi(t)$. In Dirac’s words, this is a unit ket vector $|\psi\rangle$, which is an element of the **Hilbert space** $L^2(\psi)$ with a coordinate basis $(q^i)$. The state ket–vector $|\psi(t)\rangle$ is subject to action of the Hermitian operators, obtained by the procedure of quantization of classical biodynamic quantities, and whose real eigenvalues are being measured.

**Quantum superposition** is a generalization of the algebraic principle of linear combination of vectors. The Hilbert space has a set of states $|\varphi_i\rangle$ (where the index $i$ runs over the degrees–of–freedom of the system) that form a basis and the most general state of such a system can be written as $|\psi\rangle = \sum_i c_i |\varphi_i\rangle$. The system is said to be in a state $|\psi(t)\rangle$, describing...
the motion of the de Broglie waves (named after Nobel Laureate, Prince Louis V.P.R. de Broglie), which is a linear superposition of the basis states $|\varphi_i\rangle$ with weighting coefficients $c_i$ that can in general be complex. At the microscopic or quantum level, the state of the system is described by the wave function $|\psi\rangle$, which in general appears as a linear superposition of all basis states. This can be interpreted as the system being in all these states at once. The coefficients $c_i$ are called the probability amplitudes and $|c_i|^2$ gives the probability that $|\psi\rangle$ will collapse into state $|\varphi\rangle$ when it decoheres (interacts with the environment). By simple normalization we have the constraint that $\sum_i |c_i|^2 = 1$. This emphasizes the fact that the wavefunction describes a real, physical system, which must be in one of its allowable classical states and therefore by summing over all the possibilities, weighted by their corresponding probabilities, one must get unity. In other words, we have the normalization condition for the psi–function, determining the unit length of the state ket–vector

$$\langle\psi(t)|\psi(t)\rangle = \int \psi^* \psi \, dV = \int |\psi|^2 \, dV = 1,$$

where $\psi^* = \langle\psi(t)|$ denotes the bra vector, the complex-conjugate to the ket $\psi = |\psi(t)\rangle$, and $\langle\psi(t)|\psi(t)\rangle$ is their scalar product, i.e., Dirac bracket. For this reason the scene of quantum mechanics is the functional space of square-integrable complex psi–functions, i.e., the Hilbert space $L^2(\psi)$.

When the system is in the state $|\psi(t)\rangle$, the average value $\langle f \rangle$ of any physical observable $f$ is equal to

$$\langle f \rangle = \langle\psi(t)|\hat{f}|\psi(t)\rangle,$$

where $\hat{f}$ is the Hermitian operator corresponding to $f$.

A quantum system is coherent if it is in a linear superposition of its basis states. If a measurement is performed on the system and this means that the system must somehow interact with its environment, the superposition is destroyed and the system is observed to be in only one basis state, as required classically. This process is called reduction or collapse of the wavefunction or simply decoherence and is governed by the form of the wavefunction $|\psi\rangle$.

Entanglement on the other hand, is a purely quantum phenomenon and has no classical analogue. It accounts for the ability of quantum systems to exhibit correlations in counterintuitive ‘action–at–a–distance’ ways. Entanglement is what makes all the difference in the operation of quantum computers versus classical ones. Entanglement gives ‘special powers’ to
quantum computers because it gives quantum states the potential to exhibit and maintain correlations that cannot be accounted for classically. Correlations between bits are what make information encoding possible in classical computers. For instance, we can require two bits to have the same value thus encoding a relationship. If we are to subsequently change the encoded information, we must change the correlated bits in tandem by explicitly accessing each bit. Since quantum bits exist as superpositions, correlations between them also exist in superposition. When the superposition is destroyed (e.g., one qubit is measured), the correct correlations are instantaneously ‘communicated’ between the qubits and this communication allows many qubits to be accessed at once, preserving their correlations, something that is absolutely impossible classically.

More precisely, the first quantization is a linear representation of all classical dynamical variables (like coordinate, momentum, energy, or angular momentum) by linear Hermitian operators acting on the associated Hilbert state–space $L^2(\psi)$, which has the following properties [Dirac (1949)].

1. **Linearity:**

   $$\alpha f + \beta g \rightarrow \alpha \hat{f} + \beta \hat{g},$$

   for all constants $\alpha, \beta \in \mathbb{C}$;

2. A ‘dynamical’ variable, equal to unity everywhere in the phase–space, corresponds to unit operator: $1 \rightarrow \hat{I}$; and

3. **Classical Poisson brackets**

   $$\{f, g\} = \frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i}$$

   quantize to the corresponding commutators

   $$\{f, g\} \rightarrow -i\hbar [\hat{f}, \hat{g}], \quad [\hat{f}, \hat{g}] = \hat{f}\hat{g} - \hat{g}\hat{f}.$$
transformations in the state–space by direct application to a state, i.e.,

\[ \dot{u} = \{u, f\}, \quad \partial_t |\psi\rangle = \frac{i}{\hbar} \hat{f} |\psi\rangle. \tag{2.4} \]

Exponent of anti–Hermitian operator is unitary. Due to this fact, transformations, generated by Hermitian operators

\[ \hat{U} = \exp \left( \frac{i\hat{f}t}{\hbar} \right) \]

are unitary. They are motions – scalar product preserving transformations in the Hilbert state–space \( L^2(\psi) \). For this property \( i \) is needed in (2.4).

Due to property (2), the transformations, generated by classical variables and quantum operators, have the same algebra.

For example, the quantization of energy \( E \) gives:

\[ E \rightarrow \hat{E} = i\hbar \partial_t. \]

The relations between operators must be similar to the relations between the relevant physical quantities observed in classical mechanics.

For example, the quantization of the classical equation \( E = H \), where

\[ H = H(p_i, q_i) = T + U \]
denotes the Hamilton’s function of the total system energy (the sum of the kinetic energy \( T \) and potential energy \( U \)), gives the Schrödinger equation of motion of the state ket–vector \( |\psi(t)\rangle \) in the Hilbert state–space \( L^2(\psi) \)

\[ i\hbar \partial_t |\psi(t)\rangle = \hat{H} |\psi(t)\rangle. \]

In the simplest case of a single particle in the potential field \( U \), the operator of the total system energy – Hamiltonian is given by:

\[ \hat{H} = -\frac{\hbar^2}{2m} \nabla^2 + U, \]

where \( m \) denotes the mass of the particle and \( \nabla \) is the classical gradient operator. So the first term on the r.h.s denotes the kinetic energy of the system, and therefore the momentum operator must be given by:

\[ \hat{p} = -i\hbar \nabla. \]

Now, for each pair of states \( |\varphi\rangle, |\psi\rangle \) their scalar product \( \langle \varphi | \psi \rangle \) is introduced, which is [Nikitin (1995)].
Quantum Leap

(1) Linear (for right multiplier):
\[
⟨\varphi|α_1ψ_1 + α_2ψ_2⟩ = α_1⟨\varphi|ψ_1⟩ + α_2⟨\varphi|ψ_2⟩;
\]

(2) In transposition transforms to complex conjugated:
\[
⟨ψ|φ⟩ = ⟨φ|ψ⟩;
\]
this implies that it is ‘anti–linear’ for left multiplier:
\[
⟨α_1φ_1 + α_2φ_2|ψ⟩ = \overline{α_1}⟨φ_1|ψ⟩ + \overline{α_2}⟨φ_2|ψ⟩;
\]

(3) Additionally it is often required, that the scalar product should be positively defined:
for all \(|ψ⟩\), \(⟨ψ|ψ⟩ ≥ 0\) and \(⟨ψ|ψ⟩ = 0\) iff \(|ψ⟩ = 0\).

Complex conjugation of classical variables is represented as Hermitian conjugation of operators. We remind some definitions:

- two operators \(\hat{f}, \hat{f}^+\) are called Hermitian conjugated (or adjoint), if
\[
⟨φ|\hat{f}ψ⟩ = ⟨\hat{f}^+φ|ψ⟩ \quad (\text{for all } φ, ψ).
\]
This scalar product is also denoted by \(⟨φ|\hat{f}|ψ⟩\) and called a matrix element of an operator.

- operator is Hermitian (self-adjoint) if \(\hat{f}^+ = \hat{f}\) and anti-Hermitian if \(\hat{f}^+ = -\hat{f}\);
- operator is unitary, if \(\hat{U}^+ = \hat{U}^{-1}\); such operators preserve the scalar product:
\[
⟨\hat{U}φ|\hat{U}ψ⟩ = ⟨φ|\hat{U}^+\hat{U}ψ⟩ = ⟨φ|ψ⟩.
\]

Real classical variables should be represented by Hermitian operators; complex conjugated classical variables \((a, \overline{a})\) correspond to Hermitian conjugated operators \((\hat{a}, \hat{a}^+)\).

Multiplication of a state by complex numbers does not change the state physically.

Any Hermitian operator in Hilbert space has only real eigenvalues:
\[
\hat{f}|ψ_i⟩ = f_i|ψ_i⟩, \quad (\text{for all } f_i ∈ \mathbb{R}).
\]
Eigenvectors \(|ψ_i⟩\) form complete orthonormal basis (eigenvectors with different eigenvalues are automatically orthogonal; in the case of multiple
If the two operators $\hat{f}$ and $\hat{g}$ commute, i.e., $\commutator{\hat{f}, \hat{g}} = 0$ (see Heisenberg picture below), then the corresponding quantities can simultaneously have definite values. If the two operators do not commute, i.e., $\commutator{\hat{f}, \hat{g}} \neq 0$, the quantities corresponding to these operators cannot have definite values simultaneously, i.e., the general Heisenberg uncertainty relation is valid:

$$(\Delta \hat{f})^2 \cdot (\Delta \hat{g})^2 \geq \frac{\hbar}{4} |\commutator{\hat{f}, \hat{g}}|,$$

where $\Delta$ denotes the deviation of an individual measurement from the mean value of the distribution. The well-known particular cases are ordinary uncertainty relations for coordinate–momentum ($q - p$), and energy–time ($E - t$):

$$\Delta q \cdot \Delta p \geq \frac{\hbar}{2} \quad \text{and} \quad \Delta E \cdot \Delta t \geq \frac{\hbar}{2}.$$

For example, the rules of commutation, analogous to the classical ones written by the Poisson’s brackets, are postulated for canonically–conjugate coordinate and momentum operators:

$$[\hat{q}^i, \hat{q}^j] = 0, \quad [\hat{p}_i, \hat{p}_j] = 0, \quad [\hat{q}^i, \hat{p}_j] = i\hbar \delta^i_j \hat{I},$$

where $\delta^i_j$ is the Cronecker’s symbol. By applying the commutation rules to the system Hamiltonian $\hat{H} = \hat{H}(\hat{p}_i, \hat{q}^i)$, the quantum Hamilton’s equations are obtained:

$$\frac{d(\hat{p}_i)}{dt} = -\frac{\partial \hat{H}}{\partial \hat{q}^i}, \quad \text{and} \quad \frac{d(\hat{q}^i)}{dt} = \frac{\partial \hat{H}}{\partial \hat{p}_i}.$$

A quantum state can be observed either in the coordinate $q$–representation, or in the momentum $p$–representation. In the $q$–representation, operators of coordinate and momentum have respective forms: $\hat{q} = q$, and $\hat{p}_q = -i\hbar \frac{\partial}{\partial q}$, while in the $p$–representation, they have respective forms: $\hat{q} = i\hbar \frac{\partial}{\partial p}$, and $\hat{p}_q = p_q$. The forms of the state vector $|\psi(t)\rangle$ in these two representations are mathematically related by a Fourier–transform pair (within the Planck constant).
2.1.3 Quantum Pictures

In the $q$-representation the quantum state is usually determined, i.e., the first quantization is performed, in one of the three quantum pictures (see e.g. [Dirac (1949)]):

(1) Schrödinger picture,

(2) Heisenberg picture, and

(3) Dirac interaction picture.

These three pictures mutually differ in the time-dependence, i.e., time-evolution of the state vector $|\psi(t)\rangle$ and the Hilbert coordinate basis $(q^i)$ together with the system operators.

1. In the Schrödinger (S) picture, under the action of the evolution operator $\hat{S}(t)$ the state-vector $|\psi(t)\rangle$ rotates:

$$|\psi(t)\rangle = \hat{S}(t) |\psi(0)\rangle,$$

and the coordinate basis $(q^i)$ is fixed, so the operators are constant in time:

$$\hat{F}(t) = \hat{F}(0) = \hat{F},$$

and the system evolution is determined by the Schrödinger wave equation:

$$i\hbar \frac{\partial}{\partial t} |\psi_S(t)\rangle = \hat{H}_S |\psi_S(t)\rangle.$$

If the Hamiltonian does not explicitly depend on time, $\hat{H}(t) = \hat{H}$, which is the case with the absence of variables of macroscopic fields, the state vector $|\psi(t)\rangle$ can be presented in the form:

$$|\psi(t)\rangle = \exp \left(-i\frac{E}{\hbar} t\right) |\psi\rangle,$$

satisfying the time-independent Schrödinger equation

$$\hat{H} |\psi\rangle = E |\psi\rangle,$$

which gives the eigenvalues $E_m$ and eigenfunctions $|\psi_m\rangle$ of the Hamiltonian $\hat{H}$.

2. In the Heisenberg (H) picture, under the action of the evolution operator $\hat{S}(t)$, the coordinate basis $(q^i)$ rotates, so the operators of physical variables evolve in time by the similarity transformation:

$$\hat{F}(t) = \hat{S}^{-1}(t) \hat{F}(0) \hat{S}(t),$$
while the state vector \( |\psi(t)\rangle \) is constant in time:
\[
|\psi(t)\rangle = |\psi(0)\rangle = |\psi\rangle,
\]
and the system evolution is determined by the Heisenberg equation of motion:
\[
\text{i} \hbar \partial_t \hat{F}^H(t) = [\hat{F}^H(t), \hat{H}^H(t)],
\]
where \( \hat{F}(t) \) denotes arbitrary Hermitian operator of the system, while the commutator, i.e., Poisson quantum bracket, is given by:
\[
[\hat{F}(t), \hat{H}(t)] = \hat{F}(t) \hat{H}(t) - \hat{H}(t) \hat{F}(t) = iK.
\]
In both Schrödinger and Heisenberg picture the evolution operator \( \hat{S}(t) \) itself is determined by the Schrödinger–like equation:
\[
\text{i} \hbar \partial_t \hat{S}(t) = \hat{H} \hat{S}(t),
\]
with the initial condition \( \hat{S}(0) = \hat{I} \). It determines the Lie group of transformations of the Hilbert space \( L^2(\psi) \) in itself, the Hamiltonian of the system being the generator of the group.

3. In the Dirac interaction (I) picture both the state vector \( |\psi(t)\rangle \) and coordinate basis \( (q^i) \) rotate; therefore the system evolution is determined by both the Schrödinger wave equation and the Heisenberg equation of motion:
\[
\text{i} \hbar \partial_t |\psi^I(t)\rangle = \hat{H}^I |\psi^I(t)\rangle, \quad \text{and} \quad \text{i} \hbar \partial_t \hat{F}^I(t) = [\hat{F}^I(t), \hat{H}^O(t)].
\]
Here: \( \hat{H} = \hat{H}^0 + \hat{H}^I \), where \( \hat{H}^0 \) corresponds to the Hamiltonian of the free fields and \( \hat{H}^I \) corresponds to the Hamiltonian of the interaction.

Finally, we can show that the stationary Schrödinger equation
\[
\hat{H} \psi = \hat{E} \psi
\]
can be obtained from the condition for the minimum of the quantum action:
\[
\delta S = 0.
\]
The quantum action is usually defined by the integral:
\[
S = \langle \psi(t) | \hat{H} |\psi(t)\rangle = \int \psi^* \hat{H} \psi \ dV,
\]
with the additional normalization condition for the unit–probability of the psi–function:

$$\langle \psi(t) | \psi(t) \rangle = \int \psi^* \psi \, dV = 1.$$  

When the functions $\psi$ and $\psi^*$ are considered to be formally independent and only one of them, say $\psi^*$ is varied, we can write the condition for an extreme of the action:

$$\delta S = \int \delta \psi^* \hat{H} \psi \, dV - E \int \delta \psi^* \psi \, dV = \int \delta \psi^* (\hat{H} \psi - E \psi) \, dV = 0,$$

where $E$ is a Lagrangian multiplier. Owing to the arbitrariness of $\delta \psi^*$, the Schrödinger equation $\hat{H} \psi - E \psi = 0$ must hold.

### 2.1.4 Spectrum of a Quantum Operator

To recapitulate, each state of a system is represented by a state vector $|\psi\rangle$ with a unit–norm, $\langle \psi | \psi \rangle = 1$, in a complex Hilbert space $\mathcal{H}$, and vice versa. Each system observable is represented by a Hermitian operator $\hat{A}$ in a Hilbert space $\mathcal{H}$, and vice versa. A Hermitian operator $\hat{A}$ in a Hilbert space $\mathcal{H}$ has its domain $\mathcal{D}_\hat{A} \subset \mathcal{H}$ which must be dense in $\mathcal{H}$, and for any two state vectors $|\psi\rangle, |\varphi\rangle \in \mathcal{D}_\hat{A}$ holds $\langle \hat{A} \psi | \varphi \rangle = \langle \psi | \hat{A} \varphi \rangle$ (see, e.g., [Messiah (2000)]).

#### Discrete Spectrum

A Hermitian operator $\hat{A}$ in a finite–dimensional Hilbert space $\mathcal{H}_d$ has a discrete spectrum $\{a_i, \, a \in \mathbb{R}, \, i \in \mathbb{N}\}$, defined as a set of discrete eigenvalues $a_i$, for which the characteristic equation

$$\hat{A}|\psi\rangle = a|\psi\rangle$$  \hspace{1cm} (2.5)

has the solution eigenvectors $|\psi_a\rangle \neq 0 \in \mathcal{D}_\hat{A} \subset \mathcal{H}_d$. For each particular eigenvalue $a$ of a Hermitian operator $\hat{A}$ there is a corresponding discrete characteristic projector $\hat{\pi}_a = |\psi_a\rangle \langle \psi_a|$ (i.e., the projector to the eigensubspace of $\hat{A}$ composed of all discrete eigenvectors $|\psi_a\rangle$ corresponding to $a$).

Now, the discrete spectral form of a Hermitian operator $\hat{A}$ is defined as

$$\hat{A} = a_i \hat{\pi}_i = \sum_i a_i |i\rangle \langle i|, \quad \text{for all } i \in \mathbb{N}$$  \hspace{1cm} (2.6)

where $a_i$ are different eigenvalues and $\hat{\pi}_i$ are the corresponding projectors subject to

$$\sum_i \hat{\pi}_i = \hat{I}, \quad \hat{\pi}_i \hat{\pi}_j = \delta_{ij} \hat{\pi}_j,$$
where $\hat{I}$ is identity operator in $\mathcal{H}_d$.

A Hermitian operator $\hat{A}$ defines, with its characteristic projectors $\hat{\pi}_i$, the spectral measure of any interval on the real axis $\mathbb{R}$; for example, for a closed interval $[a, b] \subset \mathbb{R}$ holds

$$\hat{\pi}_{[a,b]}(\hat{A}) = \sum_{a_i \in [a, b]} \hat{\pi}_i,$$

and analogously for other intervals, $(a, b], [a, b), (a, b) \subset \mathbb{R}$; if $a_i \in [a, b] = \emptyset$ then $\hat{\pi}_{[a,b]}(\hat{A}) = 0$, by definition.

Now, let us suppose that we measure an observable $\hat{A}$ of a system in state $|\psi\rangle$. The probability $P$ to get a result within the a priori given interval $[a, b] \subset \mathbb{R}$ is given by its spectral measure

$$P([a, b], \hat{A}, \psi) = \langle \psi | \hat{\pi}_{[a,b]}(\hat{A}) | \psi \rangle.$$

As a consequence, the probability to get a discrete eigenvalue $a_i$ as a result of measurement of an observable $\hat{A}$ equals its expected value

$$P(a_i, \hat{A}, \psi) = \langle \psi | \hat{\pi}_i | \psi \rangle = \langle \hat{\pi}_i \rangle,$$

where $\langle \hat{B} \rangle$ in general denotes the average value of an operator $\hat{B}$. Also, the probability to get a result $a$ which is not a discrete eigenvalue of an observable $\hat{A}$ in a state $|\psi\rangle$ equals zero.

**Continuous Spectrum.** A Hermitian operator $\hat{A}$ in an infinite-dimensional Hilbert space $\mathcal{H}_c$ (the so-called rigged Hilbert space) has both a discrete spectrum $\{a_i, a \in \mathbb{R}, i \in \mathbb{N}\}$ and a continuous spectrum $[c, d] \subset \mathbb{R}$. In other words, $\hat{A}$ has both a discrete sub-basis $\{|i\rangle : i \in \mathbb{N}\}$ and a continuous sub-basis $\{|s\rangle : s \in [c, d] \subset \mathbb{R}\}$. In this case $s$ is called the continuous eigenvalue of $\hat{A}$. The corresponding characteristic equation is

$$\hat{A}|\psi\rangle = s|\psi\rangle.$$

Equation (2.9) has the solution eigenvectors $|\psi_s \rangle \neq 0 \in \mathcal{D}_A \subset \mathcal{H}_c$, given by the Lebesgue integral

$$|\psi_s \rangle = \int_a^b \psi(s) |s\rangle \, ds, \quad c \leq a < b \leq d,$$

where $\psi(s) = \langle s | \psi \rangle$ are continuous, square integrable Fourier coefficients,

$$\int_a^b |\psi(s)|^2 \, ds < +\infty,$$
Quantum Leap

while the continuous eigenvectors \( |\psi_s\rangle \) are orthonormal,

\[
\psi(t) = \langle t | \psi_s \rangle = \int_c^d \psi(s) \delta(s-t) \, ds,
\]

(2.10)
i.e., normed on the Dirac \( \delta \)–function, with

\[
\langle t | s \rangle = \delta(s-t), \quad s, t \in [c, d].
\]

The corresponding continuous projectors \( \hat{\pi}_{[a,b]}^c(\hat{A}) \) are defined as Lebesgue integrals

\[
\hat{\pi}_{[a,b]}^c(\hat{A}) = \int_a^b |s\rangle \, ds \langle s| = |s\rangle \langle s|, \quad -c \leq a < b \leq d.
\]

(2.11)

In this case, projecting any vector \( |\psi\rangle \in \mathcal{H}_c \) using \( \hat{\pi}_{[a,b]}^c(\hat{A}) \) is given by

\[
\hat{\pi}_{[a,b]}^c(\hat{A}) |\psi\rangle = \left( \int_a^b |s\rangle \, ds \langle s| \right) |\psi\rangle = \int_a^b \psi(s) \, |s\rangle \, ds.
\]

Now, the continuous spectral form of a Hermitian operator \( \hat{A} \) is defined as

\[
\hat{A} = \int_c^d |s\rangle \, s \, ds \langle s|.
\]

(2.13)

**Total Spectrum.** The *total Hilbert state–space* of the system is equal to the orthogonal sum of its discrete and continuous subspaces,

\[
\mathcal{H} = \mathcal{H}_d \oplus \mathcal{H}_c.
\]

(2.12)

The corresponding discrete and continuous projectors are mutually complementary,

\[
\hat{\pi}_{a}(\hat{A}) + \hat{\pi}_{[c,d]}^c(\hat{A}) = \hat{I}.
\]

Using the closure property

\[
\sum_i |i\rangle \langle i| + \int_a^b |s\rangle \, ds \langle s| = \hat{I},
\]

the total spectral form of a Hermitian operator \( \hat{A} \in \mathcal{H} \) is given by

\[
\hat{A} = \sum_i a_i |i\rangle \langle i| + \int_c^d |s\rangle \, s \, ds \langle s|,
\]

(2.13)
while an arbitrary vector $|\psi\rangle \in \mathcal{H}$ is equal to

$$|\psi\rangle = \sum_i \psi_i |i\rangle + \int_c^d \psi(s) |s\rangle \, ds.$$  

Here, $\psi_i = \langle i|\psi\rangle$ are discrete Fourier coefficients, while $\psi(s) = \langle s|\psi\rangle$ are continuous, square integrable, Fourier coefficients,

$$\int_a^b |\psi(s)|^2 \, ds < +\infty.$$

Using both discrete and continuous Fourier coefficients, $\psi_i$ and $\psi(s)$, the total inner product of $\mathcal{H}$ is defined as

$$\langle \varphi|\psi\rangle = \varphi_1 \psi_1 + \int_c^d \varphi(s) \psi(s) \, ds,$$  

while the norm is

$$\langle \psi|\psi\rangle = \bar{\psi}_1 \psi_1 + \int_c^d \bar{\psi}(s) \psi(s) \, ds.$$

The total spectral measure is now given as

$$\hat{\pi}_{[a,b]}(\hat{A}) = \sum_i \hat{\pi}_i + \int_a^b |s\rangle \, ds \langle s|,$$

so the probability $P$ to get a measurement result within the a priori given interval $[a, b] \in \mathbb{R} \subset \mathcal{H}$ is given by

$$P([a,b], \hat{A}, \psi) = \sum_i \langle \psi|\hat{\pi}_i|\psi\rangle + \int_a^b |\psi(s)|^2 \, ds,$$  

where $|\psi(s)|^2 = \langle \psi|s\rangle \langle s|\psi\rangle$ is called the probability density. From this the expectation value of an observable $\hat{A}$ is equal to

$$\langle \hat{A} \rangle = \sum_i a_i \langle \psi|\hat{\pi}_i|\psi\rangle + \int_a^b s |\psi(s)|^2 \, ds = \langle \psi|\hat{A}|\psi\rangle.$$

### 2.1.5 General Representation Model

In quantum mechanics the total spectral form of the complete observable is given by relation (2.13). We can split this total spectral form into:
Quantum Leap

(1) Pure discrete spectral form,
\[ \hat{A} = \sum_i a_i |i\rangle \langle i|, \]
with its discrete eigenbasis \( \{|i\rangle : i \in \mathbb{N} \} \), which is orthonormal \( \langle ii|j\rangle = \delta_{ij} \) and closed \( \sum_i |i\rangle \langle i| = \hat{I} \); and

(2) Pure continuous spectral form,
\[ \hat{B} = \int_c^d \! |s\rangle \langle s| ds, \]
with its continuous eigenbasis \( \{|s\rangle : s \in [c,d] \subset \mathbb{R} \} \), which is orthonormal \( \langle ss|tt\rangle = \delta(s-t) \) and closed \( \int_c^d \! |s\rangle \langle s| ds = \hat{I} \).

The completeness property of each basis means that any vector \( |\psi\rangle \in \mathcal{H} \) can be expanded/developed along the components of the corresponding basis. In case of the discrete basis we have
\[ |\psi\rangle = \hat{I} |\psi\rangle = \sum_i |i\rangle \langle i| \psi_i = \sum_i \psi_i |i\rangle, \]
with discrete Fourier coefficients of the development \( \psi_i = \langle i| \psi \rangle \).

In case of the continuous basis we have
\[ |\psi\rangle = \hat{I} |\psi\rangle = \int_c^d \! |s\rangle \langle s| \psi \rangle = \int_c^d \! \psi(s) \langle s| ds \]
with continuous Fourier coefficients of the two development \( \psi(s) = \langle s| \psi \rangle \), which are square integrable, \( \int_c^d \! |\psi(s)|^2 ds < +\infty \).

2.1.6 Direct Product Space
Let \( \mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_n \) and \( \mathcal{H} \) be \( n+1 \) given Hilbert spaces such that dimension of \( \mathcal{H} \) equals the product of dimensions of \( \mathcal{H}_i \), \( (i = 1, \ldots, n \text{ in this section}) \). We say that the composite Hilbert space \( \mathcal{H} \) is defined as a direct product of the factor Hilbert spaces \( \mathcal{H}_i \) and write
\[ \mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_n \]
if there exists a one–to–one mapping of the set of all uncorrelated vectors \( \{|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_n\rangle\} \), \( |\psi_i\rangle \in \mathcal{H}_i \), with zero inner product (i.e., \( \langle \psi_i|\psi_j\rangle = 0 \), for \( i \neq j \)) – onto their direct product \( |\psi_1\rangle \times |\psi_2\rangle \times \cdots \times |\psi_n\rangle \), so that the following conditions are satisfied:
1. Linearity per each factor:
\[
\left( \sum_{j_1=1}^{J_1} b_{j_1} \left| \psi_{j_1} \right> \right) \times \left( \sum_{j_2=1}^{J_2} b_{j_2} \left| \psi_{j_2} \right> \right) \times \ldots \times \left( \sum_{j_n=1}^{J_n} b_{j_n} \left| \psi_{j_n} \right> \right) = \\
\sum_{j_1=1}^{J_1} \sum_{j_2=1}^{J_2} \ldots \sum_{j_n=1}^{J_n} b_{j_1} b_{j_2} \ldots b_{j_n} \left| \psi_{j_1} \right> \times \left| \psi_{j_2} \right> \times \ldots \times \left| \psi_{j_n} \right>.
\]

2. Multiplicativity of scalar products of uncorrelated vectors \( \left| \psi_i \right>, \left| \varphi_i \right> \in \mathcal{H}_i \):
\[
\left( \left| \psi_1 \right> \times \left| \psi_2 \right> \times \ldots \times \left| \psi_n \right> , \left| \varphi_1 \right> \times \left| \varphi_2 \right> \times \ldots \times \left| \varphi_n \right> \right) = \left< \psi_1 | \varphi_1 \right> \times \left< \psi_2 | \varphi_2 \right> \times \ldots \times \left< \psi_n | \varphi_n \right>.
\]

3. Uncorrelated vectors generate the whole composite space \( \mathcal{H} \), which means that in a general case a vector in \( \mathcal{H} \) equals the limit of linear combinations of uncorrelated vectors, i.e.,
\[
\left| \psi \right> = \lim_{K \to \infty} \sum_{k=1}^{K} b_k \left| \psi_{k_1} \right> \times \left| \psi_{k_2} \right> \times \ldots \times \left| \psi_{k_n} \right>.
\]

Let \( \{ \left| k_i \right> \} \) represent arbitrary bases in the factor spaces \( \mathcal{H}_i \). They induce the basis \( \{ \left| k_1 \right> \times \left| k_2 \right> \times \ldots \times \left| k_n \right> \} \) in the composite space \( \mathcal{H} \).

Let \( \hat{A}_i \) be arbitrary operators (either all linear or all antilinear) in the factor spaces \( \mathcal{H}_i \). Their direct product, \( \hat{A}_1 \otimes \hat{A}_2 \otimes \ldots \otimes \hat{A}_n \) acts on the uncorrelated vectors
\[
\left( \hat{A}_1 \otimes \hat{A}_2 \otimes \ldots \otimes \hat{A}_n \right) \left( \left| \psi_1 \right> \times \left| \psi_2 \right> \times \ldots \times \left| \psi_n \right> \right) = \left( \hat{A}_1 \left| \psi_1 \right> \right) \times \left( \hat{A}_2 \left| \psi_2 \right> \right) \times \ldots \times \left( \hat{A}_n \left| \psi_n \right> \right).
\]

### 2.1.7 State-Space for n Quantum Particles

Classical state-space for the system of \( n \) particles is its 6VD phase–space \( \mathcal{P} \), including all position and momentum vectors, \( \mathbf{r}_i = (x, y, z)_i \) and \( \mathbf{p}_i = (p_x, p_y, p_z)_i \), respectively, for \( i = 1, \ldots, n \).

The quantization is performed as a linear representation of the real Lie algebra \( \mathcal{L}_P \) of the phase–space \( \mathcal{P} \), defined by the Poisson bracket \( \{ A, B \} \) of classical variables \( A, B \) – into the corresponding real Lie algebra \( \mathcal{L}_H \) of the Hilbert space \( \mathcal{H} \), defined by the commutator \( [\hat{A}, \hat{B}] \) of skew–Hermitian operators \( \hat{A}, \hat{B} \).
We start with the Hilbert space $\mathcal{H}_x$ for a single 1D quantum particle, which is composed of all vectors $|\psi_x\rangle$ of the form

$$|\psi_x\rangle = \int_{-\infty}^{+\infty} \psi(x) |x\rangle \, dx,$$

where $\psi(x) = \langle x|\psi \rangle$ are square integrable Fourier coefficients,

$$\int_{-\infty}^{+\infty} |\psi(x)|^2 \, dx < +\infty.$$

The position and momentum Hermitian operators, $\hat{x}$ and $\hat{p}$, respectively, act on the vectors $|\psi_x\rangle \in \mathcal{H}_x$ in the following way:

$$\hat{x}|\psi_x\rangle = \int_{-\infty}^{+\infty} \hat{x} \psi(x) |x\rangle \, dx, \quad \int_{-\infty}^{+\infty} \left| \psi(x) \right|^2 \, dx < +\infty,$$

$$\hat{p}|\psi_x\rangle = \int_{-\infty}^{+\infty} -i\hbar \frac{\partial}{\partial x} \psi(x) |x\rangle \, dx, \quad \int_{-\infty}^{+\infty} \left| -i\hbar \frac{\partial}{\partial x} \psi(x) \right|^2 \, dx < +\infty.$$

The orbit Hilbert space $\mathcal{H}_o^1$ for a single 3D quantum particle with the full set of compatible observable $\hat{\mathbf{r}} = (\hat{x}, \hat{y}, \hat{z})$, $\hat{\mathbf{p}} = (\hat{p}_x, \hat{p}_y, \hat{p}_z)$, is defined as

$$\mathcal{H}_o^1 = \mathcal{H}_x \otimes \mathcal{H}_y \otimes \mathcal{H}_z,$$

where $\hat{\mathbf{r}}$ has the common generalized eigenvectors of the form

$$|\hat{\mathbf{r}}\rangle = |x\rangle \times |y\rangle \times |z\rangle.$$

$\mathcal{H}_o^1$ is composed of all vectors $|\psi_r\rangle$ of the form

$$|\psi_r\rangle = \int_{-\infty}^{+\infty} \psi(x,y,z) |x\rangle \times |y\rangle \times |z\rangle \, dx \, dy \, dz,$$

where $\psi(x,y,z) = \langle r|\psi \rangle$ are square integrable Fourier coefficients,

$$\int_{-\infty}^{+\infty} \left| \psi(x,y,z) \right|^2 \, dx \, dy \, dz < +\infty.$$

The position and momentum operators, $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$, respectively, act on the vectors $|\psi_r\rangle \in \mathcal{H}_o^1$ in the following way:

$$\hat{\mathbf{r}}|\psi_r\rangle = \int_{\mathcal{H}_o^1} \hat{\mathbf{r}} \psi(x,y,z) |x\rangle \times |y\rangle \times |z\rangle \, dx \, dy \, dz, \quad \int_{\mathcal{H}_o^1} \left| \psi(x,y,z) \right|^2 \, dx \, dy \, dz < +\infty,$$

$$\hat{\mathbf{p}}|\psi_r\rangle = \int_{\mathcal{H}_o^1} -i\hbar \frac{\partial}{\partial \mathbf{r}} \psi(x,y,z) |x\rangle \times |y\rangle \times |z\rangle \, dx \, dy \, dz, \quad \int_{\mathcal{H}_o^1} \left| -i\hbar \frac{\partial}{\partial \mathbf{r}} \psi(x,y,z) \right|^2 \, dx \, dy \, dz < +\infty.$$
Now, if we have a system of $n$ 3D particles, let $\mathcal{H}_i^o$ denote the orbit Hilbert space of the $i$th particle. Then the composite orbit state–space $\mathcal{H}_n^o$ of the whole system is defined as a direct product

$$\mathcal{H}_n^o = \mathcal{H}_1^o \otimes \mathcal{H}_2^o \otimes \ldots \otimes \mathcal{H}_n^o.$$}

$\mathcal{H}_n^o$ is composed of all vectors

$$|\psi_n^o\rangle = \int_{\mathcal{H}_n^o} \psi(r_1, r_2, \ldots, r_n) |r_1\rangle \times |r_2\rangle \times \ldots \times |r_n\rangle \, dr_1 dr_2 \ldots dr_n$$

where $\psi(r_1, r_2, \ldots, r_n) = \langle r_1, r_2, \ldots, r_n|\psi_n^o\rangle$ are square integrable Fourier coefficients

$$\int_{\mathcal{H}_n^o} |\psi(r_1, r_2, \ldots, r_n)|^2 \, dr_1 dr_2 \ldots dr_n < +\infty.$$

The position and momentum operators $\hat{\mathbf{r}}_i$ and $\hat{\mathbf{p}}_i$ act on the vectors $|\psi_n^o\rangle \in \mathcal{H}_n^o$ in the following way:

$$\hat{\mathbf{r}}_i |\psi_n^o\rangle = \int_{\mathcal{H}_n^o} \{\hat{\mathbf{r}}_i\} \psi(r_1, r_2, \ldots, r_n) |r_1\rangle \times |r_2\rangle \times \ldots \times |r_n\rangle \, dr_1 dr_2 \ldots dr_n,$$

$$\hat{\mathbf{p}}_i |\psi_n^o\rangle = \int_{\mathcal{H}_n^o} \{-i\hbar \frac{\partial}{\partial \mathbf{r}_i}\} \psi(r_1, r_2, \ldots, r_n) |r_1\rangle \times |r_2\rangle \times \ldots \times |r_n\rangle \, dr_1 dr_2 \ldots dr_n,$$

with the square integrable Fourier coefficients

$$\int_{\mathcal{H}_n^o} |\{\hat{\mathbf{r}}_i\} \psi(r_1, r_2, \ldots, r_n)|^2 \, dr_1 dr_2 \ldots dr_n < +\infty,$$

$$\int_{\mathcal{H}_n^o} \left|\{-i\hbar \frac{\partial}{\partial \mathbf{r}_i}\} \psi(r_1, r_2, \ldots, r_n)\right|^2 \, dr_1 dr_2 \ldots dr_n < +\infty,$$

respectively. In general, any set of vector Hermitian operators $\{\hat{\mathbf{A}}_i\}$ corresponding to all the particles, act on the vectors $|\psi_n^o\rangle \in \mathcal{H}_n^o$ in the following way:

$$\hat{\mathbf{A}}_i |\psi_n^o\rangle = \int_{\mathcal{H}_n^o} \{\hat{\mathbf{A}}_i\} \psi(r_1, r_2, \ldots, r_n) |r_1\rangle \times |r_2\rangle \times \ldots \times |r_n\rangle \, dr_1 dr_2 \ldots dr_n,$$

with the square integrable Fourier coefficients

$$\int_{\mathcal{H}_n^o} \left|\{\hat{\mathbf{A}}_i\} \psi(r_1, r_2, \ldots, r_n)\right|^2 \, dr_1 dr_2 \ldots dr_n < +\infty.$$
2.2 Relativistic Quantum Mechanics and Electrodynamics

2.2.1 Difficulties of the Relativistic Quantum Mechanics

The theory outlined above is not in agreement with the Einstein’s restricted Principle of relativity, as is at once evident from the special role played by the time \( t \). Thus, while it works very well in the non-relativistic region of low velocities, where it appears to be in complete agreement with experiment, it can be considered only as an approximation, and one must face the task of extending it to make it conform to restricted relativity. One should be prepared for possible further alterations being needed in basic physical concepts, and hence one should follow the route of first setting up the mathematical formalism and then seeking its physical interpretation.

Setting up the mathematical formalism is a fairly straightforward matter. One must first put classical Newtonian mechanics into relativistic Hamiltonian form. One must take into account that the various particles comprising the dynamical system interact through the medium of the electromagnetic field, and one must use Lorentz’s equations of motion for them, including the damping terms which express the reaction of radiation. This is done in subsection 2.2.4 below, where, with the help of the Dirac’s electrodynamic action principle, the equations of motion are obtained in the Hamiltonian form (2.62) with the Hamiltonians \( F_i \), one for each particle, given by (2.61). This Hamiltonian formulation may now be made into a quantum theory by following rules which have become standardized from the non-relativistic quantum mechanics. The resulting formalism appears to be quite satisfactory mathematically, but when one proceeds to consider its physical interpretation one meets with serious difficulties.[Dirac (1926c); Dirac (1926e); Dirac (1932); Chandra (1948)].

Take an elementary example, that of a free particle without spin, moving in the absence of any field. The classical Hamiltonian for this system is the left-hand side of the equation

\[
p_0^2 - p_1^2 - p_2^2 - p_3^2 - m^2 = 0,
\]

(2.16)

where \( p_0 \) is the energy and \( p_1, p_2, p_3 \) the momentum of the particle, the velocity of light being taken as unity. Passing over to quantum theory by the standard rules, one gets from this Hamiltonian the so-called Klein–
Gordon equation

\[(\hbar^2 \Box + m^2)\psi = 0,\]  \hspace{1cm} (2.17)

where \(\Box\) is the Dalambertian wave operator,

\[\Box \equiv \frac{\partial^2}{\partial x_0^2} - \frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} - \frac{\partial^2}{\partial x_3^2}.\]

The wave function \(\psi\) here is a scalar, involving the coordinates \(x_1, x_2, x_3\) and the time \(t = x_0\) on the same footing, and so it is suitable for a relativistic theory.

If one now tries to use the old interpretation that \(|\psi|^2\) is the probability per unit volume of the particle being in the neighborhood of the point \(x = x_1, x_2, x_3\) at the time \(x_0\), one immediately gets into conflict with relativity, since this probability ought to transform under Lorentz transformations like the time–component of a 4–vector, while \(|\psi|^2\) is a scalar. Also the conservation law for total probability would no longer hold, the usual proof of it failing on account of the wave equation (2.17) not being linear in \(\partial_{x_0} \equiv \partial/\partial x_0\).

An important step forward was taken by \[\text{Gordon} (1926)\] and \[\text{Klein} (1927)\], who proposed that instead of \(|\psi|^2\) one should use the expression

\[\frac{1}{4\pi}\left[\psi \partial_{x_0} \bar{\psi} - \bar{\psi} \partial_{x_0} \psi\right],\]

(2.18)

where \(\bar{\psi} = \bar{\psi}(x_0, x_1, x_2, x_3)\) is the complex–conjugate wave \(\psi–\)function.

The expression (2.18) is the time component of a 4–vector. Further, it is easily verified that the divergence of this 4–vector vanishes, which gives the conservation law in relativistic form. Thus, (2.18) is evidently the correct mathematical form to use.

However, this form leads to trouble on the physical side, since, although it is real, it is not positive definite like \(|\psi|^2\). Its employment would result in one having at times a negative probability for the particle being in a certain place.

This is not the only physical difficulty. Let us consider the energy and momentum of the particle, and take for simplicity a state for which these variables have definite values. The corresponding wave \(\psi–\)function will be of the form of plane waves,

\[\psi = \exp[-i(p_0 x_0 - p_1 x_1 - p_2 x_2 - p_3 x_3)/\hbar].\]
In order that the wave equation (2.17) may be satisfied, the energy and momentum values \(p_0, p_1, p_2, p_3\) here must satisfy the classical equation (2.16). This equation allows of negative values for the energy \(p_0\) as well as positive ones and is, in fact, symmetrical between positive and negative energies. The negative energies occur also in the classical theory, but do not then cause trouble, since a particle started off in a positive-energy state can never make a transition to a negative-energy one. In the quantum theory, however, such transitions are possible and do in general take place under the action of perturbing forces [Dirac (1926c); Dirac (1926e); Dirac (1932)].

The wave \(\psi\)-function may be transformed to the momentum and energy variables. The Klein–Gordon expression (2.18) then goes over into

\[
|\psi(p_0, p_1, p_2, p_3)|^2 p_0^{-1} dp_1 dp_2 dp_3, \quad (2.19)
\]
as the probability of the momentum having a value within the small domain \(dp_1 dp_2 dp_3\) about the value \(p_1, p_2, p_3\), with the energy having the value \(p_0\), which must be connected with \(p_1, p_2, p_3\) by (2.16). The weight factor \(p_0^{-1}\) appears in (2.19) and makes it Lorentz invariant, since \(\psi(p)\) is a scalar (it is defined in terms of \(\psi(x)\) to make it so), and the differential element \(p_0^{-1} dp_1 dp_2 dp_3\) is also Lorentz invariant. This weight factor may be positive or negative, and makes the probability positive or negative accordingly. Thus the two undesirable things, negative energy and negative probability, always occur together.

Let us pass on to another simple example, that of a free particle with spin half a quantum. The wave equation is of the same form (2.17) as before, but the wave \(\psi\)-function is no longer a scalar. It must have two components, or four if there is a field present, and the way they transform under Lorentz transformations is given by the general connection between the theory of angular momentum in quantum mechanics and group theory. The expression \(\sum |\psi(x)|^2\), summed for the components of \(\psi\), turns out to be the time component of a 4-vector, and further the divergence of this 4-vector vanishes. Thus it is satisfactory to use this expression as the probability per unit volume of the particle being at any place at any time. One does not now have any negative probabilities in the theory. However, the negative energies remain, as in the case of no spin.

We may go on and consider particles of higher spin. The general result is that there are always states of negative energy as well as those of positive energy. For particles whose spin is an integral number of quanta, the
negative–energy states occur with a negative probability and the positive–energy ones with a positive probability, while for particles whose spin is a half–odd integral number of quanta, all states occur with a positive probability \cite{Dirac1926e,Dirac1932}.

Negative energies and probabilities should not be considered as nonsense. They are well–defined concepts mathematically, like a negative sum of money, since the equations which express the important properties of energies and probabilities can still be used when they are negative. Thus negative energies and probabilities should be considered simply as things which do not appear in experimental results. The physical interpretation of relativistic quantum mechanics that one gets by a natural development of the non–relativistic theory involves these things and is thus in contradiction with experiment. We therefore have to consider ways of modifying or supplementing this interpretation.

2.2.2 Particles of Half–Odd Integral Spin

Let us first consider particles with a half–odd integral spin, for which there is only the negative–energy difficulty to be removed. The chief particle of this kind for which a relativistic theory is needed is the electron, with spin half a quantum. Now electrons, and also, it is believed, all particles with a half–odd integral spin, satisfy the Pauli’s Exclusion Principle, according to which not more than one of them can be in any quantum state.\footnote{This principle is obtained in quantum mechanics from the requirement that wave functions shall be antisymmetric in all the particles.} With this principle there are only two alternatives for a state, either it is unoccupied or it is occupied by one particle, and a symmetry appears with respect to these two alternatives.

Dirac proposed a way of dealing with the negative–energy difficulty for electrons, based on a theory in which nearly all their negative–energy states are occupied (see \cite{Dirac1936}). An unoccupied negative–energy state now appears as a ‘hole’ in the distribution of occupied negative–energy states and thus has a deficiency of negative energy, i.e., a positive energy. From the wave equation one finds that a hole moves in the way one would expect a positively charged electron to move. It becomes reasonable to identify the holes with the recently discovered positrons, and thus to get an interpretation of the theory involving positrons together with electrons. An electron jumping from a positive– to a negative–energy state in the theory is now interpreted as an annihilation of an electron and a positron, and
one jumping from a negative– to a positive–energy state as a creation of an
electron and a positron.

The theory involves an infinite density of electrons everywhere. It be-
comes necessary to assume that the distribution of electrons for which all
positive–energy states are unoccupied and all negative–energy states occu-
pied, what one may call the vacuum distribution, as it corresponds to the
absence of all electrons and positrons in the interpretation, is completely
unobservable. Only departures from this distribution are observable and
contribute to the electric density and current which give rise to electromag-
netic field in accordance with Maxwell’s equations.

The above theory does provide a way out from the negative–energy
difficulty, but it is not altogether satisfactory. The infinite number
electrons that it involves requires one to deal with wave functions of
very great complexity and leads to such complicated mathematics that
one cannot solve even the simplest problems accurately, but must re-
sort to crude and unreliable approximations. Such a theory is a most
inconvenient one to have to work with, and on general philosophical
grounds one feels that it must be wrong [Dirac (1926c); Dirac (1926e);
Dirac (1936)].

Let us see whether one can modify the theory so as to make it possible
to work out simple examples accurately, while retaining the basic idea of
identifying unoccupied negative–energy states with positrons. The simple
calculations that one can make involve simple wave functions, referring
to only one or two electrons, and thus referring to nearly all the negative–
ergy states being unoccupied. The calculations therefore apply to a world
almost saturated with positrons, i.e., having nearly every quantum state for
a positron occupied. Such a world, of course, differs very much from the
actual world. One can now calculate the probability of any kind of collision
process occurring in this hypothetical world (in so far as electrons and
positrons are concerned). One can deduce the probability coefficient for
the process, i.e., the probability per unit number of incident particles or
per unit intensity of the beam of incident particles, for each of the various
kinds of incident particle taking part in the process. For this purpose one
must use the laws of statistical mechanics, which tell how the probability
of a collision process depends on the number of incident particles, paying
due attention to the modified form of these laws arising from the Pauli’s
exclusion principle.

Let us now assume that probability coefficients so calculated for the
hypothetical world are the same as those of the actual world. This single
assumption provides a general physical interpretation for the formalism, enabling one to calculate collision probabilities in the actual world. It does not provide a complete physical theory, since it enables one to calculate only those experimental results that are reducible to collision probabilities, and some branches of physics, e.g., the structure of solids, do not seem to be so reducible. However, collision probabilities are the things for which a relativistic theory is at present most needed, and one may hope in the future to find ways of extending the scope of the theory to make it include the whole of physics.

Comparing the new theory with the old, one may say that the new assumption, identifying collision probability coefficients in the actual world with those in a certain hypothetical world, replaces the old assumption about the non–observability of the vacuum distribution of negative–energy electrons. The approximations needed for working out simple examples in the old theory are equivalent in their mathematical effect to making the new assumption; e.g., these approximations include the neglect of the Coulomb interaction between electron and positron in the calculation of the probability of pair creation and annihilation, and this interaction cannot appear in the new theory, since the calculation there is concerned with a one–electron system. Thus the new theory may be considered as a precise formulation of the old theory together with some general approximations needed for applying it.

The new theory for dealing with the negative–energy states of the electron may be applied to any kind of elementary particle with spin half a quantum, and probably also to particles with other half–odd integral spin values, provided, of course, they satisfy Pauli’s exclusion principle. It may thus be applied to protons and neutrons. It requires for each particle the possibility of existence of an antiparticle of the opposite charge, if the original particle is charged. If the original particle is uncharged, one can arrange for the antiparticle to be identical with the original [Dirac (1926c); Dirac (1926e); Dirac (1936)].

2.2.3 Particles of Integral Spin

Most of the elementary particles of physics have half–odd integral spin, but there is the important exception of the photon (or, light–quantum), with spin one quantum, and there is the cosmic–ray particle, the meson, also probably with spin one quantum. All these kinds of particle, it is believed, satisfy the Bose–Einstein statistics, a statistics which allows any
number of particles to be in the same quantum state with the same \textit{a priori} probability\footnote{This statistics is obtained in quantum mechanics from the requirement that wave functions shall be symmetric in all the particles.}. For these kinds of particles the previous method of dealing with the negative–energy states is therefore no longer applicable, and there is the further difficulty of the negative probabilities.

When dealing with particles satisfying the Bose–Einstein statistics, it is useful to consider the operators corresponding to the absorption of a particle from a given state or the emission into a given state. These operators can be treated as dynamical variables, although they do not have any analogues in classical mechanics. If one works out their equations of motion and transformation equations, one finds a remarkable correspondence. The absorption operators from a set of independent states have the same equations of motion and transformation equations as the wave $\psi$–function representing a single particle, and similarly for the emission operators and the conjugate complex wave $\bar{\psi}$–function. Thus one can pass from a one–particle theory to a many–particle theory by making the $\psi$ and $\bar{\psi}$ describing the one particle into absorption and emission operators (or annihilation and creation operators), which must satisfy the appropriate commutation relations. Such a passage is called \textit{second quantization}.

One can get over the difficulties of negative energies and negative probabilities for Bose–Einstein particles by abandoning the attempt to get a satisfactory theory of a single particle and passing on to consider the problem of many particles, using a method given by Pauli and Weisskopf \cite{Pauli1934} for electrons having no spin and satisfying the Bose–Einstein statistics\footnote{Such electrons are not known experimentally, but there is no known theoretical reason why they should not exist.}. The method of Pauli and Wiesskopf is to work entirely with positive–energy states. The operators of absorption from and emission into negative–energy states, arising in the application of second quantization to the one–electron theory, are replaced by the operators of emission into and absorption from positive–energy states of electrons with the opposite charge, respectively. This replacement does not disturb the laws of conservation of charge, energy and momentum. The resulting theory involves spinless electrons of both kinds of charge together, and leads to pair creation and annihilation, as with ordinary electrons and positrons \cite{Dirac1926c,Dirac1926e}.

The method of Pauli and Wiesskopf may be applied in a degenerate form to photons and leads to the quantum electrodynamics of Heisenberg and...
Pauli [Heisenberg and Pauli (1929a); Heisenberg and Pauli (1929b)]. To take into account that photons have no charge, one must start with a one–
particle theory in which the wave functions are real, so that $\bar{\psi} = \psi$. The part of the wave $\psi$–function referring to positive–energy states is then made into the absorption operators from positive–energy states, and the part referring to negative–energy states into the emission operators into positive energy states. The resulting scheme of operators, involving only positive energy photon states, may then be put into correspondence with classical electrodynamics, according to the usual laws governing the correspondence between quantum and classical theory.

It would seem that in this way the difficulties of negative energies and probabilities for Bose–Einstein particles can be overcome, but a new difficulty appears. When one tries to solve the wave equation (or the wave equations if there are several particles with their respective Hamiltonians) one gets divergent integrals in the solution, of the form, in the case of photons,

$$\int_0^\infty f(v)dv, \quad f(v) \sim v^n \text{ for large } v,$$

(2.20)

$v$ being the frequency of a photon. The values 1, 0 and $-1$ for $n$ are the chief ones occurring in simple examples. Thus the wave equation really has no solutions and the method fails [Dirac (1926c); Dirac (1926e)].

Dirac had made a detailed study of the divergent integrals occurring in quantum electrodynamics and had shown [Dirac (1936)] with even values of $n$ can be eliminated by introducing into the equations a certain limiting process, which one can justify by showing that a corresponding limiting process is needed in classical electrodynamics to get the equations of motion into Hamiltonian form (which appears according to the Dirac’s electrodynamic action principle, see subsection 2.2.4 below). The divergent integrals with odd values of $n$ remain, however, and indicate something more fundamentally wrong with the theory.

Divergent integrals are a general feature of quantum field theories, and it has usually been supposed that they should be avoided by altering the forces or the laws of interaction between the elementary particles at small distances, so as to get the integrals cut off for some high value of $v$. However, one can easily see that this is wrong, in the case of electrodynamics at any rate, by referring to the corresponding classical theory. The wave $\psi$–function should have its analogue in the solution of the Hamilton–Jacobi equation, in accordance with equation (2.2), but already when one tries
Quantum Leap

to solve the Hamilton–Jacobi equation of classical electrodynamics corresponding to the wave equation of Heisenberg and Pauli’s quantum electrodynamics, one meets with divergent integrals. Now the classical equations of motion concerned, namely, Lorentz’s equations including radiation damping, have definite solutions when treated by straightforward methods and if, on trying to get these solutions by a Hamilton–Jacobi method, one meets with divergent integrals, it means simply that the Hamilton–Jacobi method is an unsuitable one, and not that one should try to alter the physical laws of interaction to get the integrals to converge. The correspondence between the quantum and classical theories is so close that one can infer that the corresponding divergent integrals in the quantum theory must also be due to an unsuitable mathematical method.

The appearance of divergent integrals with odd \( n \)–values in Heisenberg and Pauli’s form of quantum electrodynamics may be ascribed to the asymmetrical treatment of positive– and negative–energy photon states. If instead of using Pauli and Weisskopf’s method one keeps to plain second quantization, one can build up a form of quantum electrodynamics symmetrical between positive– and negative–energy photon states \( \text{Dirac (1926e)} \) \( \text{Dirac (1936)} \). The new theory leads to similar equations as the old one, but with integrals of the type

\[
\int_{-\infty}^{\infty} f(v)dv, \tag{2.21}
\]

instead of \( (2.20) \), and since \( f(v) \) is always a rational algebraic function, and it is reasonable on physical grounds to approach the upper and lower limits of integration in \( (2.21) \) at the same rate, the divergencies with odd \( n \)–values all cancel out.

Dirac had shown that the new form of quantum electrodynamics also corresponds to classical electrodynamics in accordance with the usual laws, with the exception that operators corresponding to real dynamical variables in the classical theory are no longer always self-adjoint. This exception is not important, as it rather stands apart from the general mathematical connection between quantum and classical theory. The Hamilton–Jacobi equation corresponding to the wave equation of the new quantum electrodynamics differs from that of the old one only through being expressed in terms of a different set of coordinates, but the new Hamilton–Jacobi equation can be solved without divergent integrals and is connected with a satisfactory action principle \( \text{Dirac (1932)} \) \( \text{Dirac (1926e)} \) \( \text{Dirac (1936)} \). Thus the correspondence with classical theory of the new
form of quantum electrodynamics is more far-reaching than that of the old form, which provides a strong reason for preferring the new form. It now becomes necessary to find some new physical interpretation to avoid the difficulties of negative energies and probabilities.

Let us consider in more detail the relation between the two forms of quantum electrodynamics. In either form the electromagnetic potentials $A$ at two points $x'$ and $x''$ must satisfy the commutation relations

$$[A_\mu(x'), A_\nu(x'')] = g_{\mu\nu} \Delta(x' - x''),$$

obtained from analogy with the classical theory, $\Delta$ being the four-dimensional Lorentz-invariant function introduced by Jordan and Pauli (1928), which has a singularity on the light-cone and vanishes everywhere else. In the quantum electrodynamics of Heisenberg and Pauli the $A$’s are operators referring to the absorption and emission of photons into positive energy states. Let us call such operators $A^1$. One could introduce a similar set of operators referring to the absorption and emission of photons into negative-energy states. Let us call these operators $A^2$. They satisfy the same commutation relations (2.22) and commute with the $A^1$’s. One can now introduce a third set of operators

$$A^3 = \frac{\sqrt{2}}{2} (A^1 + A^2),$$

which operate on wave functions referring to photons in both positive- and negative-energy states, and which satisfy the same commutation relations (2.22). The use of this third set gives the new form of quantum electrodynamics arising from plain second quantization.

The three sets of $A$’s may be expressed in terms of their Fourier components as [Dirac (1926e)] [Dirac (1932)] [Dirac (1936)]

$$A^1(x) = \int \left[ R_k e^{i(kx)} + \bar{R}_k e^{-i(kx)} \right] k_0^{-1} dk_1 dk_2 dk_3, \quad (k_0 = \sqrt{k^2_1 + k^2_2 + k^2_3}),$$

where $\int$ denotes the triple integral, $R_k$ is the emission operator and $\bar{R}_k$ is the absorption operator,

$$A^1(x) = \int \left[ R_k e^{i(kx)} + \bar{R}_k e^{-i(kx)} \right] k_0^{-1} dk_1 dk_2 dk_3, \quad (k_0 = -\sqrt{k^2_1 + k^2_2 + k^2_3}),$$

(2.23)
\[ A^3(x) = \frac{\sqrt{2}}{2} \sum_{k_0=\pm} \int \left[ R_k e^{i(k,x)} + \bar{R}_k e^{-i(k,x)} \right] k_0^{-1} dk_1 dk_2 dk_3. \] (2.25)

Since the three sets of \( A \)'s all satisfy the same commutation relations, they must correspond merely to three different representations of the same dynamical variables, and the passage from one to another must be a transformation of the linear type usual in quantum mechanics. Therefore, after obtaining the divergency–free solution of the wave equation in the representation corresponding to \( A^3 \), one could apply a transformation to get the solution in the \( A^1 \) representation. However, the transformation would then introduce the same divergent integrals as appear with the direct solution of the wave equation in the \( A^1 \) representation, so one would not get any further this way [Dirac (1936)].

In working with the \( A^3 \) representation one has redundant dynamical variables. It is as though, in dealing with a system of one degree of freedom with the variables \( q, p \), one decided to treat it as a system of two degrees–of–freedom by putting
\[ q = \frac{\sqrt{2}}{2} (q_1 + q_2) \quad \text{and} \quad p = \frac{\sqrt{2}}{2} (p_1 + p_2). \]

This would be quite a correct procedure, but would introduce an unnecessary complication. In the case of quantum electrodynamics, the complication is a necessary one, to avoid the divergent integrals. Let us put
\[ B(x) = \frac{\sqrt{2}}{2} [A^1(x) - A^2(x)]. \] (2.26)

Then the \( B \)'s commute with the \( A^3 \)'s, and thus with all the dynamical variables appearing in the Hamiltonian, so they are the redundant variables.

To determine the significance of redundant variables in quantum mechanics one may consider a general case, and work in a representation which separates the redundant variables from the non–redundant ones. One then sees immediately that a solution of the wave equation corresponds in general, not to a single state, but to a set of states with a certain probability for each, what in the classical theory is called a Gibbs ensemble. The probabilities of the various states depend on the weights attached to the various eigenvalues of the redundant variables in the wave \( \psi \)–function, these weights being arbitrary, depending on the weight factor in the representation used. If one works in a representation which does not separate
the redundant and non-redundant variables, as is the case in quantum electrodynamics with the representation corresponding to the use of $A^3$, the general result that wave functions represent Gibbs ensembles and not single states must still be valid. Thus one can conclude that there are no solutions of the wave equation of quantum electrodynamics representing single states, but only solutions representing Gibbs ensembles. The problem remains of interpreting the negative energies and probabilities occurring with these Gibbs ensembles.

For any $x$, $B(x)$ commutes with the Hamiltonian and is a constant of the motion. We may give it any value we like, subject to not contradicting the commutation relations. Instead of $B(x)$ it is more convenient to work with the potential field, $F(x)$ say, obtained from $B(x)$ by changing the sign of all the Fourier components containing $e^{ik_0x_0}$ with negative values of $k_0$. From (2.26), (2.23) and (2.24), we have

$$B(x) = \frac{\sqrt{2}}{2} \sum_{k_0=\pm} \int \left[ R_k \psi (k,x) - \bar{R}_k \psi^* (k,x) \right] k_0^{-1} dk_1 dk_2 dk_3. \quad (2.27)$$

Let us now take $B$ equal to the initial value of $A^3$, a proceeding which does not contradict the commutation relations since its consequences are self-consistent. Then for the initial wave function we have

$$[B(x) - A^3(x)] \psi = 0,$$

or, from (2.25) and (2.27),

$$\bar{R}_k \psi = 0, \quad (2.28)$$

with $k_0$ either positive or negative. Thus any absorption operator applied to the initial wave function gives the result zero, which means that the corresponding state is one with no photons present.

The following natural interpretation for the wave function at some later time now appears. That part of it corresponding to no photons present may be supposed to give (through the square of its modulus) the probability of no change having taken place in the field of photons; that part corresponding to one positive-energy photon present may be supposed to give the probability of a photon having been emitted; that corresponding to one negative-energy photon present may be supposed to give the probability of a photon having been absorbed; and so on for the parts corresponding to two or more photons present. The various parts of the wave function which referred to the existence of positive- and negative-energy photons in
the old interpretation now refer to the emissions and absorptions of photons. This disposes of the negative–energy difficulty in a satisfactory way, conforming to the laws of conservation of energy and momentum. It is possible only because of the redundant variables enabling one to arrange that the initial wave $\psi$–function shall correspond in its entirety to no emissions or absorptions having taken place.

The interpretation is not yet complete, because the theory at present would give a negative probability for a process involving the absorption of a photon, or the absorption of any odd number of photons. To find the origin of these negative probabilities, one must study the probability distribution of the photons initially present in the Gibbs ensemble, which one can do by transforming to the representation corresponding to the $A^1$ potentials. It is true that one cannot apply this transformation to a solution of the wave equation without getting divergent integrals, as has already been mentioned, but one can apply it to the initial wave $\psi$–function, which is of a specially simple form in the photon variables. In [Dirac (1932); Dirac (1926e); Dirac (1936)] it is found that the probability of there being $n$ photons initially in any photon state is $P_n = \pm 2$, according to whether $n$ is even or odd. Strictly, to make $\sum_{n=0}^{\infty} P_n$ converge to the limit unity, one must consider $P_n$ as a limit,

$$P_n = 2(\epsilon - 1)^n, \quad (2.29)$$

with $\epsilon$ a small positive quantity tending to zero.

Probabilities 2 and $-2$ are, clearly, not physically understandable, but one can use them mathematically in accordance with the rules for working with a Gibbs ensemble. One can suppose a hypothetical mathematical world with the initial probability distribution (2.29) for the photons, and one can work out the probabilities of radiative transition processes occurring in this world. One can deduce the corresponding probability coefficients, i.e., the probabilities per unit intensity of each beam of incident radiation concerned, by using Einstein’s laws of radiation. For example, for a process involving the absorption of a photon, if the probability coefficient is $B$, the probability of the process is

$$\sum_{n=0}^{\infty} nP_nB = \frac{1}{2}B, \quad (2.30)$$

and for a process involving the emission of a photon, if the probability
coefficient is $A$, the probability of the process is

$$\sum_{n=0}^{\infty} (n + 1) P_n A = \frac{1}{2} A. \quad (2.31)$$

Now the probability of an absorption process, as calculated from the theory, is negative, and that for an emission process is positive, so that, equating these calculated probabilities to (2.30) and (2.31) respectively, one obtains positive values for both $B$ and $A$. Generally, it is easily verified that any radiative transition probability coefficient obtained by this method is positive.

It now becomes reasonable to assume that these probability coefficients obtained for a hypothetical world are the same as those of the actual world. One gets in this way a general physical interpretation for the quantum theory of photons. When applied to elementary examples, it gives the same results as Heisenberg and Pauli’s quantum electrodynamics with neglect of the divergent integrals, since the extra factor $\sqrt{2}/2$ occurring in the matrix elements of the present theory owing to the $\sqrt{2}/2$ in the right–hand side of (2.25) compensates the factor $1/2$ in the right–hand side of (2.30) or (2.31). The present general method of physical interpretation is probably applicable to any kind of particle with an integral spin [Dirac (1932); Dirac (1926); Chandra (1948)].

Therefore, it appears that, whether one is dealing with particles of integral spin or of half-odd integral spin, one is led to a similar conclusion, namely, that the mathematical methods at present in use in quantum mechanics are capable of direct interpretation only in terms of a hypothetical world differing very markedly from the actual one. These mathematical methods can be made into a physical theory by the assumption that results about collision processes are the same for the hypothetical world as the actual one. One thus gets back to Heisenberg’s view about physical theory, that all it does is to provide a consistent means of calculating experimental results. The limited kind of description of Nature which Schrödinger’s method provides in the non–relativistic case is possible relativistically only for the hypothetical world, and even then is rather more indefinite (e.g., the principle of superposition of states no longer applies), because of the need to use a Gibbs ensemble for describing the photon distribution.

To have a description of Nature is philosophically satisfying, though not logically necessary, and it is somewhat strange that the attempt to get such a description should meet with a partial success, namely, in the non–
relativistic domain, but yet should fail completely in the later development. It seems to suggest that the present mathematical methods are not final. Any improvement in them would have to be of a very drastic character, because the source of all the trouble, the symmetry between positive and negative energies arising from the association of energies with the Fourier components of functions of the time, is a fundamental feature of them [Dirac (1932); Dirac (1926b); Dirac (1936); Chandra (1948)].

2.2.4 Dirac’s Electrodynamics Action Principle

There are various forms which the action principle of classical electrodynamics may take, but most of them involve awkward conditions concerning the singularities of the field where the charged particles are situated and are not suitable for a subsequent passage to quantum mechanics.

Fokker [Fokker (1929)] set up a form of action principle which does not refer to the singularities of the field and which appears to be the best starting point for getting a quantum theory. Fokker’s action integral may conveniently be written with the help of the $\delta$–function as

$$S = S_1 + S_2,$$

where

$$S_1 = \sum_i m_i \int ds_i,$$

and

$$S_2 = \sum_i \sum_{j \neq i} e_i e_j \int \int \delta(z_i - z_j)^2 (v_i, v_j) ds_i ds_j. \tag{2.33}$$

Here, the scalar product notation is used as

$$(a, b) = a^\mu b_\mu = a_0 b_0 - a_1 b_1 - a_2 b_2 - a_3 b_3,$$

and $m_i$ and $e_i$ are the mass and charge of the $i$th particle, the 4–vector $z_i$ gives the four coordinates of the point on the world–line of the $i$th particle whose proper–time is $s_i$, and $v_i$ is the velocity 4–vector of the $i$th particle satisfying

$$v_i = \frac{dz_i}{ds_i}, \tag{2.34}$$

and

$$v_i^2 = 1. \tag{2.35}$$

The integrals in (2.32)–(2.33) are taken along the world–lines of the particles, and the occurrence of the $\delta$–function $\delta(z_i - z_j)^2$ in $S_2$ ensures that the only values for $z_i$ and $z_j$ contributing to the double integral are those for which
Elements of Quantum Mechanics

\((z_i - z_j)^2 = 0\), which means that each of the points \(z_i, z_j\) is on the past or future light–cone from the other.

The action integral as it stands is not a general one covering all possible states of motion. To make it general one must, as has been pointed out by the Dirac (1938), add to it a term of the form

\[
S_3 = \sum_i e_i \int M_\mu(z_i) v_i^\mu ds_i.
\]  

(2.36)

The 4–vector potential \(M_\mu(x)\) may be left for the present an arbitrary function of the field point \(x\).

For the purpose of deducing the equations of motion, one may take the limits of integration in the various integrals to be \(-\infty\) and \(\infty\), as was done by Fokker, but in order to introduce momenta and get the equations into Hamiltonian form one must take finite limits. Let us therefore suppose that each \(s_i\) goes from \(s_0^i\) to \(s'_i\), and let the corresponding \(z_i\) and \(v_i\) be \(z_0^i, z'_i\) and \(v_0^i, v'_i\). It is desirable to restrict the initial values \(s_0^i\) so that the points \(z_0^i\) all lie outside each other’s light–cones, and similarly with the final values \(s'_i\). Thus

\[
(z_0^i - z_0^j)^2 < 0, \quad (z'_i - z'_j)^2 < 0, \quad (i \neq j).
\]  

(2.37)

Now, before making variations in \(S\), one should replace \(S_1\), by

\[
S'_1 = \sum_i m_i \int \sqrt{v_i^2 ds_i},
\]  

so as to make \(S\) homogeneous of degree zero in the differential elements \(ds_i, v_i\) counting as being of degree \(-1\) [Dirac (1926e)]. The expression for \(S\) is then valid with \(s_i\) any parameter on the world–line of the \(i\)th particle, so that \(v_i\) defined by (2.34) does not necessarily satisfy (2.35).

Let us now make variations \(\partial z_i(s_i)\) in the world–lines of the particles, \(\partial M(x)\) in the field function \(M(x)\), and \(D s'_i\) in the final values of the \(s_i\), so that the end–points of the world–lines are changed by

\[
D z'_i = \partial z'_i + v'_i D s'_i,
\]  

(2.39)

\(\partial z'_i\) being written for \(\partial z_i(s'_i)\). The initial values of the \(s_i\) and the initial points of the world–lines we suppose for simplicity to be fixed, since variations in them would give rise to terms of the same form as those arising from variations in the final values and would not lead to anything new.
Varying $S'_1$ given by (2.38) and using (2.35), after the variation process, one gets with the help of (2.39),

$$
S'_1 = \sum_i m_i \left[ - \int_{s^0}^{s^1} \left( \frac{dv_i}{ds_i}, \partial z_i \right) ds_i + (v'_i, Dz'_i) \right].
$$

(2.40)

To get the variation in $S_2$ given by (2.33) one may, owing to the symmetry between $i$ and $j$ in the double sum, vary only quantities involving $i$ and multiply by 2. The result is [Dirac (1936)]

$$
\partial S_2 = \sum_i \sum_{j \neq i} e_i e_j \int_{s^0}^{s^1} \int_{s^0}^{s^1} \left[ \partial \frac{\delta(z_i - z_j)^2}{\partial z_i} (v_i, v_j) \right] ds_i ds_j
$$

$$
- \frac{d}{ds_i} [\delta(z_i - z_j)^2 v_j] \partial z_i ds_i + \int_{s^0}^{s^1} \delta(z'_i - z_j)^2 (v'_j, Dz'_i) ds_j.
$$

(2.41)

Finally, in varying $S_3$ given by (2.36), one has to take into account that the total variation in $M$ at a point $z_i(s_i)$ on the $i$th world-line, let us call it $DM(z_i)$, consists of two parts, a part $\partial M(z_i)$ arising from the variation in the function $M(x)$ and equal to the value of $\partial M(x)$ at the point $x = z_i$, and a part arising from the variation in $z_i$, equal to $\partial M/\partial x$, at the point $x = z_i$ multiplied into $\partial z_i$; thus

$$
DM(z_i) = \partial M(z_i) + (\partial M/\partial x) z_i, \partial z_i.
$$

(2.42)

The variation in $S_3$ is now [Dirac (1936)]

$$
\partial S_3 = \sum_i e_i \left\{ \int_{s^0}^{s^1} \left[ \partial M^\mu(z_i) v_{\mu i} + \left( \frac{\partial M^\mu}{\partial x^\nu} \right)_{z_i} v_{\mu i} \partial z_{\nu i} - \frac{dM^\mu(z_i)}{ds_i} \partial z_{\mu i} \right] ds_i 
+ M^\mu(z'_i) Dz'_{\mu i} \right\}.
$$

(2.43)

The total variation in $S$ is given by the sum of the three expressions (2.40), (2.41) and (2.43).

By equating to zero the total coefficient of $\partial z_{\mu i}$, one gets the equation of motion of the $i$th particle. It is

$$
- m_i \frac{dv_i}{ds_i} + e_i \sum_{j \neq i} e_j \int_{s^0}^{s^1} \left[ \frac{\partial \delta(z_i - z_j)^2}{\partial z_i} (v_i, v_j) - \frac{d}{ds_i} [\delta(z_i - z_j)^2 v_j] \right] ds_j
$$

$$
+ e_i \left( \frac{\partial M^\mu}{\partial x^\nu} \right)_{z_i} v_{\mu i} - \frac{dM^\mu(z_i)}{ds_i} = 0.
$$
Introducing the field function
\[ A_\mu^i(x) = M_\mu^i(x) + \sum_{j \neq i} e_j \int_{s_0^j}^{s_j} \partial \delta(x - z_j)^2 v_\mu^j ds_j, \]  \hfill (2.44)
the above equation of motion may be written
\[ m_i \frac{dv_\mu^i}{ds_i} = e_i \left[ \left( \frac{\partial A_\mu^i}{\partial x_\nu} \right)_{z^i} v_\mu^i - \frac{dA_\nu^i(z_i)}{ds_i} \right] = e_i \left[ \frac{\partial A_\mu^i}{\partial x_\nu} - \frac{\partial A^i_\nu}{\partial x_\mu} \right]_{z^i} v_\mu^i. \]  \hfill (2.45)

It is the correct Lorentz equation of motion of the \( i \)th particle, provided \( A_\mu^i \) is connected with the ingoing and outgoing fields and the retarded and advanced fields of the other particles by the relation, given by Dirac (1938),
\[ A_\mu^i(x) = \frac{1}{2} [A_\mu^{in}(x) + A_\mu^{out}(x)] + \frac{1}{2} \sum_{j \neq i} e_j \left[ \int_{s_0^j}^{s_j} \delta(x - z_j)^2 v_\mu^j ds_j + \int_{s_j}^{s_0} \delta(x - z_j)^2 v_\mu^j ds_j \right]. \]  \hfill (2.46)

According to (2.44) this requires (in Dirac’s notation for integrals)
\[ M_\mu^i(x) = \frac{1}{2} [A_\mu^{in}(x) + A_\mu^{out}(x)] + \sum_{j \neq i} e_j \left[ \int_{s_0^j}^{s_j} \delta(x - z_j)^2 v_\mu^j ds_j + \int_{s_j}^{s_0} \delta(x - z_j)^2 v_\mu^j ds_j \right]. \]  \hfill (2.47)

Note that we are summing here over all values of \( j \) \footnote{It also depends on the proper–times \( s_0 \), but this does not concern us here.}, \( \text{Dirac } (1926) \). In space–time region which lies inside the future light–cone from \( z_0^i \) and inside the past light–cone from \( z_0' \). By assuming that (2.47) holds throughout space–time, one gets an expression for \( M_\mu^i(x) \) independent of \( i \), so that the equations of motion of all the particles follow from the same Fokker’s action integral.

One can now pass to the Hamiltonian formulation of the equations of motion. For each point in space–time \( x \), \( M_\mu^i(x) \) may be counted as a coordinate, depending on the proper–times \( s_0^i \) and will have a conjugate momenta, say \( K_\mu^i(x) \). These momenta, together with the particle momenta \( p_\mu^i \), are defined, as in the general theory \footnote{Dirac (1926)}, by the coefficients of \( \partial M_\mu^i(x) \) and \( Dz_\mu^i \) in the expression for \( \partial S \), so that we have
\[ \partial S = \sum_i p_\mu^i Dz_\mu^i + \int_{-\infty}^{\infty} K_\mu^i(x) \partial M_\mu^i(x) dx_0 dx_1 dx_2 dx_3, \]  \hfill (2.48)
where the integral sign denotes the quadruple space–time integral. Comparing (2.48) with the sum of (2.40), (2.41) and (2.43), one gets [Dirac (1926e); Dirac (1936)]

\[ K_\mu(x) = \sum_i e_i \int_{s_i^0}^{s_i^f} \delta(x_0 - z_{0i})\delta(x_1 - z_{1i})\delta(x_2 - z_{2i})\delta(x_3 - z_{3i}) v_\mu ds_i \]  \hspace{1cm} (2.49)

and

\[ p_\mu^i = m_i v_\mu^i + e_i [M^\mu(z_i^f) + \frac{1}{2} \sum_j e_j \int_{s_j^0}^{s_j^f} \Delta(x_j^f - z_j + \lambda)v_\mu ds_j], \]  \hspace{1cm} (2.50)

where \( \lambda \) is a small 4-vector whose direction is within the future light–cone (so that \( \lambda^2 > 0, \lambda_0 > 0 \)), \( \Delta(y) \) denotes the Jordan and Pauli (1928) \( \Delta \)–function of any 4-vector \( y \), satisfying the 4D wave equation (here \( \Box \) is the Dalambertian wave operator, \( \Box = \partial_0^2 - \partial_1^2 + \partial_2^2 + \partial_3^2 \))

\[ \Box \Delta(y) = 0 \]  \hspace{1cm} which implies \[ \Box M^\mu(y) = 0, \]

and related to the corresponding \( \delta \)–function by

\[ \Delta(y) = \pm 2\delta(y^2). \]

The momenta satisfy the Poisson bracket commutation relationships

\[ [p_{\mu i}, z_{\nu j}] = g_{\mu\nu}\delta_{ij}, \]  \hspace{1cm} (2.51)

\[ [K_\mu(x), M_\nu(x')] = g_{\mu\nu}\delta(x_0 - x'_0)\delta(x_1 - x'_1)\delta(x_2 - x'_2)\delta(x_3 - x'_3), \]  \hspace{1cm} (2.52)

so that the Poisson bracket of any two momenta or of any two coordinates vanishes. Instead of \( K_\mu(x) \) it is more convenient to work with the momentum field–function \( N_\mu(x) \) defined by [Dirac (1958); Dirac (1929)]

\[ N_\mu(x) = \frac{1}{2} \int _{-\infty}^{\infty} \Delta(x - x')K_\mu(x')dx_0 dx_1 dx_2 dx_3, \]  \hspace{1cm} (2.53)

and satisfying \[ \Box N_\mu(x) = 0. \]  \hspace{1cm} (2.54)

Instead of (2.52) one has

\[ [N_\mu(x), M_\nu(x')] = \frac{1}{2} g_{\mu\nu}\Delta(x - x'). \]  \hspace{1cm} (2.55)

From (2.53) and (2.49) one gets

\[ N_\mu(x) = \frac{1}{2} \sum_i e_i \int_{s_i^0}^{s_i^f} \Delta(x - z_i)v_\mu ds_i, \]  \hspace{1cm} (2.56)
so that (2.50) may be written

\[ p_i^\mu = m_i v_i^\mu + e_i \left[ M^\mu(z_i^') + N^\mu(z_i^' + \lambda) \right] = m_i v_i^\mu + e_i A^\mu(z_i^'), \tag{2.57} \]

where \( A^\mu(x) = M^\mu(x) + N^\mu(x + \lambda). \tag{2.58} \)

From (2.54) the potentials \( A_\mu(x) \) satisfy

\[ \Box A_\mu(x) = 0, \tag{2.59} \]

showing that they can be resolved into waves travelling with the velocity of light, and from (2.55) it follows

\[ [A_\mu(x), A_\nu(x')] = \frac{1}{2} g_{\mu\nu} \left[ \Delta(x - x' + \lambda) + \Delta(x - x' - \lambda) \right]. \tag{2.60} \]

From (2.35) and (2.57) it follows

\[ F_i \equiv \left[ p_i^\mu - e_i A_\mu(z_i^') \right]^2 - m_i^2 = 0. \tag{2.61} \]

There is one of these equations for each particle. The expressions \( F_i \) may be used as Hamiltonians to determine how any dynamical variable \( \xi \) varies with the proper–times \( s_i' \), in accordance with the equations \cite{dirac1958, dirac1926, dirac1949} \( \kappa_i \frac{d\xi}{ds_i'} = [\xi, F_i], \tag{2.62} \)

were \( \xi \) is any function of the coordinates and momenta of the particles and of the fields \( M, K, N, A \), and the \( \kappa \)'s are multiplying factors not depending on \( \xi \). Taking \( \xi = z_i^' \), one finds that

\[ \kappa_i = -2m_i, \]

to get agreement with (2.57). Taking \( \xi = p_i^\mu \) gives one back the equation of motion (2.45) with the \( \lambda \) refinement. Taking \( \xi = M_\mu(x) \), one gets from (2.58) and (2.55),

\[ \frac{M_\mu(x)}{ds_i'} = e_i v_i^\mu [M_\mu(x), A_\nu(z_i^')] = \frac{1}{2} e_i v_i^\mu \Delta(x - z_i^' - \lambda). \]

This equation of motion for the field quantities \( M_\mu(x) \) does not follow from the variation principle, as it involves only coordinates and velocities and not accelerations, and it has to be imposed as an extra condition in the variational method.
The above Hamiltonian formulation of the equations of classical electrodynamics may be taken over into the quantum theory in the usual way, by making the momenta into operators satisfying commutation relations corresponding to the Poisson bracket relations (2.51), (2.52). Equation (2.60) in the limit $\lambda \rightarrow 0$ goes over into the quantum equation (2.22). The Hamiltonians (2.61) provide the wave equations

$$F_i \psi = 0,$$

in which the wave $\psi$–function is a function of the coordinates $z_i'$ of all the particles and of the field variables $M_\mu(x)$. One can apply the theory to spinning electrons instead of spinless particles, by modifying the Hamiltonians $F_i$ in the appropriate way. For more details, see [Dirac (1958); Dirac (1926e); Dirac (1949)].

2.2.5 Dirac Equation and Formal QED in Brief

Recall (see, e.g., [Drake (2006); Ivancevic and Ivancevic (2007b)]) that in natural units, Dirac equation(s) for a particle with mass $m$ read (with $\mu = 0, 1, 2, 3$)

$$(i\gamma^\mu \partial_\mu - m)\psi = 0, \quad (i\gamma^\mu \partial_\mu - m)\bar{\psi} = 0, \quad (2.63)$$

where $\psi = \psi(x)$ is the 4–component Dirac wave spinor $\psi = \bar{\psi}(x)$ is the Dirac adjoint spinor, while $\gamma^\mu$ are $4 \times 4$ Dirac $\gamma$–matrices,

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$

$$\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

They obey the anticommutation relations

$$\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu},$$

where $g_{\mu\nu}$ is the metric tensor.

---

6The most convenient definitions for the two-spinors, like the Dirac spinor, are:

$$\phi^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \phi^2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{and} \quad \chi^1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \chi^2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$
Dirac’s $\gamma$--matrices are conventionally derived as
\[
\gamma^k = \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix}, \quad (k = 1, 2, 3)
\]
where $\sigma^k$ are Pauli $\sigma$--matrices\(^7\) (a set of $2 \times 2$ complex Hermitian and unitary matrices), defined as
\[
\sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},
\]
obeying both the commutation and anticommutation relations
\[
[\sigma_i, \sigma_j] = 2i \epsilon_{ijk} \sigma_k, \quad \{\sigma_i, \sigma_j\} = 2\delta_{ij} \cdot I,
\]
where $\epsilon_{ijk}$ is the Levi–Civita symbol, $\delta_{ij}$ is the Kronecker delta, and $I$ is the identity matrix.

The Dirac adjoint wave–spinor $\bar{\psi} = \bar{\psi}(x^\mu)$ is defined as
\[
\bar{\psi} = (\psi^\dagger) \gamma^0,
\]
where $\psi^\dagger$ is the Hermitian adjoint of $\psi$, so that we can define the probability–density current $j^\mu = j^\mu(x)$, as
\[
\bar{\psi} \gamma^\mu \psi.
\]

The Lorentz–invariant form of the Dirac equation (2.63) for an electron with a charge $e$ and mass $m_e$, moving with a 4–momentum $p_\mu$ in a classical electromagnetic field defined by its 4–potential $A_\mu = A_\mu(x)$, reads (see, e.g., Drake (2006) [Ivancevic and Ivancevic (2007b)]):
\[
\{i\gamma^\mu [p_\mu - eA_\mu] - m_e\} \psi = 0, \quad (2.64)
\]
and is called the covariant Dirac equation.

The formal QED Lagrangian (density) includes three terms,
\[
\mathcal{L}(x) = \mathcal{L}_{em}(x) + \mathcal{L}_{int}(x) + \mathcal{L}_{e-p}(x), \quad (2.65)
\]
related respectively to the free electromagnetic field $F^{\mu\nu} = F^{\mu\nu}(x)$, the electron–positron field (in the presence of the external vector potential

\(^7\)In quantum mechanics, each Pauli matrix represents an observable describing the spin of a spin $\frac{1}{2}$ particle in the three spatial directions. Also, $i\sigma_j$ are the generators of rotation acting on non-relativistic particles with spin $\frac{1}{2}$. The state of the particles are represented as two–component spinors.

In quantum information, single–qubit quantum gates are $2 \times 2$ unitary matrices. The Pauli matrices are some of the most important single–qubit operations.
$A^\mu(x)$, and the interaction field (dependent on the charge–current vector $j^\mu$). The free electromagnetic field Lagrangian in (2.65) has the standard electrodynamic form

$$L_{\text{em}}(x) = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu},$$

where the electromagnetic fields are expressible in terms of the four-potentials by

$$F^{\mu\nu} = \partial^{\mu} A^{\nu}_{\text{tot}} - \partial^{\nu} A^{\mu}_{\text{tot}}, \quad \text{with} \quad A^{\mu}_{\text{tot}}(x) = A^\mu_{\text{ext}}(x) + A^\mu(x).$$

The electron-positron field Lagrangian is given by Dirac’s equation (2.64) as

$$L_{\text{e–p}}(x) = \bar{\psi} \gamma^\mu \left[ p^\mu - e A^\mu_{\text{ext}} \right] - m_e \psi.$$

The interaction field Lagrangian

$$L_{\text{int}}(x) = -j^\mu A^\mu,$$

accounts for the interaction between the uncoupled electrons and the radiation field.

The field equations deduced from (2.65) read

$$\{ i \gamma^\mu \left[ p^\mu - e A^\mu_{\text{ext}} \right] - m_e c \} \psi = i \gamma^\mu \psi A^\mu, \quad \partial^\mu F_{\mu\nu} = j^\nu. \quad (2.66)$$

The formal QED requires the solution of the system (2.66) when $A^\mu(x)$, $\psi(x)$ and $\bar{\psi}(x)$ are quantized fields.

### 2.2.6 Lorentzian Space-Time and Gravity

Recall that in modern physics, thanks to the work of Einstein [Einstein (1916)], space and time are unified into the space-time manifold $(M, g)$, where the metric $g$ is a real–valued symmetric bilinear map $g : T_p(M) \times T_p(M) \to \mathbb{R}$ of Lorentzian signature. The latter feature gives rise to the light-cone structure of space-time, with vectors being divided into time–like, null or space–like depending on whether $g(X, X)$ is negative, vanishing or positive, respectively. The classical laws of nature are written in tensor language, and gravity is the curvature of space-time. In the theory of general relativity, gravity couples to the energy–momentum tensor of matter through the
Einstein equations (see, e.g., Ivancevic and Ivancevic (2007b))
\[
R_{ab} - \frac{1}{2} g_{ab} R = \frac{8 \pi G}{c^4} T_{ab}. \tag{2.67}
\]

The Einstein–Hilbert action functional for gravity, giving rise to equation (2.67), is diffeomorphism–invariant, and hence general relativity belongs actually to the general set of theories ruled by an infinite–dimensional invariance group (see next chapter).

In Riemannian geometry (or pseudo–Riemannian one, see Ivancevic and Ivancevic (2007b)), geodesics are curves whose tangent vector \( X \) moves by parallel transport, so that eventually
\[
\frac{dX^a}{ds} + \Gamma^a_{bc} X^b X^c = 0,
\]
where \( s \) is the affine parameter and \( \Gamma^a_{bc} \) are the affine–connection coefficients [Ivancevic and Ivancevic (2007b)]. In general relativity, time–like geodesics correspond to the trajectories of freely moving observers, while null geodesics describe the trajectories of zero–rest–mass particles Hawking and Ellis (1973). Moreover, a space–time \((M, g)\) is said to be singularity–free if all time–like and null geodesics can be extended to arbitrary values of their affine parameter. At a space–time singularity in general relativity, all laws of classical physics would break down, because one would witness very pathological events such as the sudden disappearance of freely moving observers, and one would be completely unable to predict what came out of the singularity.

2.2.7 Unification of Fundamental Interactions

The three fully established unifications of modern physics are Bavnbek et al. (2007):

(1) Maxwell: electricity and magnetism are unified into electromagnetism. All related phenomena can be described by an antisymmetric rank–two tensor field, and derived from a one–form, called the potential.

(2) Einstein: space and time are unified into the space–time manifold. Moreover, inertial and gravitational mass, conceptually different, are actually unified as well.

(3) Standard model of particle physics: electromagnetic, weak and strong forces are unified by a non–Abelian gauge theory, normally considered
in Minkowski space-time (this being the base space in fibre-bundle language [Ivancevic and Ivancevic (2007b)].

The physics community is now familiar with a picture relying upon four fundamental interactions: electromagnetic, weak, strong and gravitational. The large–scale structure of the universe, however, is ruled by gravity only. All unifications beyond Maxwell involve non–Abelian gauge groups (either Yang–Mills or diffeomorphism group [Ivancevic and Ivancevic (2007b)]). Three extreme views have been developed along the years:

1. Gravity arose first, temporally, in the very early Universe, then all other fundamental interactions.
2. Gravity might result from Quantum Field Theory (this was the Sakharov idea [Sakharov (1968)]).
3. The vacuum of particle physics is regarded as a cold quantum liquid in equilibrium. Protons, gravitons and gluons are viewed as collective excitations of this liquid.

In the following, we outline how the space–of–histories formulation provides a common ground for describing the ‘old’ and ‘new’ unifications of fundamental theories.

2.2.7.1 First Unification

Quantum field theory begins once an action functional $S$ is given, since the first and most fundamental assumption of quantum theory is that every isolated dynamical system can be described by a characteristic action functional. The Feynman approach (see next chapter) makes it necessary to consider an infinite–dimensional manifold such as the space $\Phi$ of all field histories $\phi^i$. On this space there exist (in the case of gauge theories) vector fields $Q_\alpha = Q^i_\alpha \delta_{\alpha \beta} \phi^i$, that leave the action invariant, i.e., $Q_\alpha S = 0$. The Lie brackets of these vector fields (see [Ivancevic and Ivancevic (2007b)]) lead to a classification of all gauge theories known so far.

Type–I Gauge Theories

The peculiar property of type–I gauge theories is that these Lie brackets are equal to linear combinations of the vector fields themselves, with structure constants, i.e., [DeWitt (2003)]

$$[Q_\alpha, Q_\beta] = C^\gamma_{\alpha \beta} Q_\gamma,$$ (2.68)
where $C_{\alpha\beta,i}^\gamma = 0$. The Maxwell, Yang–Mills, Einstein theories are all examples of type–I theories (this is the ‘unifying feature’). All of them, being gauge theories, need supplementary conditions, since the second functional derivative of $S$ is not an invertible operator. After imposing such conditions, the theories are ruled by a differential operator of d’Alembert type (or Laplace type, if one deals instead with Euclidean field theory), or a non-minimal operator at the very worst (for arbitrary choices of gauge parameters). For example, when Maxwell theory is quantized via path–integrals (see next chapter) in the Lorenz gauge, one deals with a gauge–fixing functional [Bavnbek et al. (2007)]

$$\Phi(A) = \nabla^b A_b,$$

and the second–order differential operator acting on the potential reads as

$$P_a^b = -\delta^b_a \Box + R_a^b + \left(1 - \frac{1}{\alpha}\right) \nabla_a \nabla^b,$$

where $\alpha$ is an arbitrary gauge parameter. The Feynman choice $\alpha = 1$ leads to the minimal operator

$$\tilde{P}_a^b = -\delta^b_a \Box + R_a^b,$$

which is the standard wave operator on vectors in curved space-time. Such operators play a leading role in the one-loop expansion of the Euclidean effective action, i.e., the quadratic order in $\hbar$ in the asymptotic expansion of the functional ruling the quantum theory with positive–definite metrics.

**Type–II Gauge Theories**

For type–II gauge theories, Lie brackets of vector fields $Q_\alpha$ are as in equation (2.68) for type–I theories, but the structure constants are promoted to structure functions. An example is given by simple supergravity (a supersymmetric gauge theory of gravity, with a symmetry relating bosonic and fermionic fields) in four space-time dimensions, with auxiliary fields [Bavnbek et al. (2007)].

**Type–III Gauge Theories**

In this case, the Lie bracket (2.68) is generalized by

$$[Q_\alpha, Q_\beta] = C_{\alpha\beta}^\gamma Q_\gamma + U_{\alpha\beta}^i S_i$$
and it therefore reduces to (2.68) only on the mass–shell, i.e., for those field configurations satisfying the Euler–Lagrange equations. An example is given by theories with gravitons and gravitinos such as Bose–Fermi supermultiplets of both simple and extended supergravity in any number of space-time dimensions, without auxiliary fields [Bavnbek et al. (2007)].

From Supergravity to General Relativity

It should be stressed that general relativity is naturally related to supersymmetry, since the requirement of gauge–invariant Rarita–Schwinger equations [Rarita and Schwinger (1941)] implies Ricci–flatness in four dimensions, which is then equivalent to vacuum Einstein equations. The Dirac operator (see Ivancevic and Ivancevic (2007b)) is more fundamental in this framework, since the mD space-time metric is entirely re-constructed from the $\gamma$–matrices, giving [Bavnbek et al. (2007)]

$$g^{ab} = \frac{1}{2m} \text{Tr}(\gamma^a \gamma^b + \gamma^b \gamma^a).$$
Chapter 3

Feynman Path Integrals

In this chapter we develop the formalism of Feynman path integrals, the essential tool in highly–nonlinear and high–dimensional dynamics.

3.1 Path Integrals: Sums Over Histories

The pivot–point of theoretical physics in the last half of a Century has been the celebrated path integral, a powerful conceptual and computational tool first conceived by Richard (Dick) Feynman, and later generalized by Ed Witten, Stephen Hawking and other pioneers of physical science. Recall that in the path–integral formalism, we first formulate the specific classical action of a new theory, and subsequently perform its quantization by means of the associated transition amplitude. This action–amplitude picture is the core structure in any new physical theory. Its virtual paths are in general neither deterministic nor smooth, although they include bundles and jets of deterministic and smooth paths, as well as Markov chains. Yet, it is essentially a (broader) geometrical dynamics, with its Riemannian and symplectic versions, among many others. At the beginning, it worked only for conservative physical systems. Today it includes also dissipative structures, as well as various sources and sinks, geometries and topologies. Its smooth part reveals all celebrated equations of physics, both classical and quantum. It is the core of modern quantum gravity and superstring theory. It is arguably the most important construct of mathematical physics. At the edge of a new millennium, if you asked a typical theoretical physicist: what will be your main research tool in the new millennium, they would most probably say: path integral. And today, we see it moving out from physics, into the realm of social sciences. Finally, since Feynman’s fairly intuitive invention of the path integral [Feynman (1951)], a lot of research has been
done to make it mathematically rigorous (see e.g., [Loo (1999); Loo (2000); Albeverio et al. (1986); Klauder (1997); Shabanov and Klauder (1998); Klauder (2000)]).

3.1.1 Intuition Behind a Path Integral

3.1.1.1 Classical Probability Concept

Recall that a random variable $X$ is defined by its distribution function $f(x)$. Its probabilistic description is based on the following rules: (i) $P(X = x_i)$ is the probability that $X = x_i$; and (ii) $P(a \leq X \leq b)$ is the probability that $X$ lies in a closed interval $[a, b]$. Its statistical description is based on: (i) $\mu_X$ or $E(X)$ is the mean or expectation of $X$; and (ii) $\sigma_X$ is the standard deviation of $X$. There are two cases of random variables: discrete and continuous, each having its own probability (and statistics) theory.

**Discrete Random Variable**

A discrete random variable $X$ has only a countable number of values $\{x_i\}$. Its distribution function $f(x_i)$ has the following properties:

$$P(X = x_i) = f(x_i), \quad f(x_i) \geq 0, \quad \sum_i f(x_i) \, dx = 1.$$  

Statistical description of $X$ is based on its discrete mean value $\mu_X$ and standard deviation $\sigma_X$, given respectively by

$$\mu_X = E(X) = \sum_i x_i f(x_i), \quad \sigma_X = \sqrt{E(X^2) - \mu_X^2}.$$  

**Continuous Random Variable**

Here $f(x)$ is a piecewise continuous function such that:

$$P(a \leq X \leq b) = \int_a^b f(x) \, dx, \quad f(x) \geq 0, \quad \int_{-\infty}^{\infty} f(x) \, dx = \int_{\mathbb{R}} f(x) \, dx = 1.$$  

Statistical description of $X$ is based on its continuous mean $\mu_X$ and standard deviation $\sigma_X$, given respectively by

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x f(x) \, dx, \quad \sigma_X = \sqrt{E(X^2) - \mu_X^2}.$$
Now, let us observe the similarity between the two descriptions. The same kind of similarity between discrete and continuous quantum spectrum stroke Dirac when he suggested the combined integral approach, that he denoted by (see, e.g., Drake (2006)): \( \int \Sigma \) meaning 'both integral and sum at once', that is, integration over the continuous spectrum and summing over the discrete spectrum.

To emphasize this similarity even further, as well as to set–up the stage for the path integral, recall the notion of a cumulative distribution function of a random variable \( X \), that is a function \( F : \mathbb{R} \to \mathbb{R} \), defined by

\[
F(a) = P(X \leq a).
\]

In particular, suppose that \( f(x) \) is the distribution function of \( X \). Then

\[
F(x) = \sum_{x_i \leq x} f(x_i), \quad \text{or} \quad F(x) = \int_{-\infty}^{x} f(t) \, dt,
\]

according to as \( x \) is a discrete or continuous random variable. In either case, \( F(a) \leq F(b) \) whenever \( a \leq b \). Also,

\[
\lim_{x \to -\infty} F(x) = 0 \quad \text{and} \quad \lim_{x \to \infty} F(x) = 1,
\]

that is, \( F(x) \) is monotonic and its limit to the left is 0 and the limit to the right is 1. Furthermore, its cumulative probability is given by

\[
P(a \leq X \leq b) = F(b) - F(a),
\]

and the Fundamental Theorem of Calculus tells us that, in the continuum case,

\[
f(x) = \frac{\partial}{\partial x} F(x).
\]

General Markov Stochastic Dynamics

Recall that Markov stochastic process is a random process characterized by a lack of memory, i.e., the statistical properties of the immediate future are uniquely determined by the present, regardless of the past [Gardiner (1985)].

For example, a random walk is an example of the Markov chain, i.e., a discrete–time Markov process, such that the motion of the system in consideration is viewed as a sequence of states, in which the transition from one state to another depends only on the preceding one, or the probability of the system being in state \( k \) depends only on the previous state \( k - 1 \).
The property of a Markov chain of prime importance in biomechanics is the existence of an invariant distribution of states: we start with an initial state $x_0$ whose absolute probability is 1. Ultimately the states should be distributed according to a specified distribution.

Between the pure deterministic dynamics, in which all DOF of the system in consideration are explicitly taken into account, leading to classical dynamical equations, for example in Hamiltonian form \( H = \frac{1}{2}p_i^2 + V(q_i) \), i.e.,

\[
\dot{q}_i = \frac{\partial p_i}{\partial q_i} H, \quad \dot{p}_i = -\frac{\partial q_i}{\partial p_i} H
\]

– and pure stochastic dynamics (Markov process), there is so-called hybrid dynamics, particularly Brownian dynamics, in which some of DOF are represented only through their stochastic influence on others. As an example, suppose a system of particles interacts with a viscous medium. Instead of specifying a detailed interaction of each particle with the particles of the viscous medium, we represent the medium as a stochastic force acting on the particle. The stochastic force reduces the dimensionally of the dynamics.

Recall that the Brownian dynamics represents the phase–space trajectories of a collection of particles that individually obey Langevin rate equations in the field of force (i.e., the particles interact with each other via some deterministic force). For a free particle, the Langevin equation reads [Gardiner (1985)]:

\[
m \dot{v} = R(t) - \beta v,
\]

where $m$ denotes the mass of the particle and $v$ its velocity. The right–hand side represent the coupling to a heat bath; the effect of the random force $R(t)$ is to heat the particle. To balance overheating (on the average), the particle is subjected to friction $\beta$. In humanoid dynamics this is performed with the Rayleigh–Van der Pol’s dissipation. Formally, the solution to the Langevin equation can be written as

\[
v(t) = v(0) \exp \left( -\frac{\beta t}{m} \right) + \frac{1}{m} \int_0^t \exp \left[ -\frac{(t-\tau)\beta}{m} \right] R(\tau) d\tau,
\]

where the integral on the right–hand side is a stochastic integral and the solution $v(t)$ is a random variable. The stochastic properties of the solution depend significantly on the stochastic properties of the random force $R(t)$. In the Brownian dynamics the random force $R(t)$ is Gaussian distributed. Then the problem boils down to finding the solution to the
Langevin stochastic differential equation with the supplementary condition (mean zero and variance)

\[ \langle R(t) \rangle = 0, \quad \langle R(t) R(0) \rangle = 2\beta k_B T \delta(t), \]

where \( \langle \cdot \rangle \) denotes the mean value, \( T \) is temperature, \( k_B \)—equipartition (i.e., uniform distribution of energy) coefficient, \( \delta(t) \)—function.

Algorithm for computer simulation of the Brownian dynamics (for a single particle) can be written as [Heermann (1990)]:

1. Assign an initial position and velocity.
2. Draw a random number from a Gaussian distribution with mean zero and variance.
3. Integrate the velocity to get \( v_{n+1} \).
4. Add the random component to the velocity.

Another approach to taking account the coupling of the system to a heat bath is to subject the particles to collisions with virtual particles [Heermann (1990)]. Such collisions are imagined to affect only momenta of the particles, hence they affect the kinetic energy and introduce fluctuations in the total energy. Each stochastic collision is assumed to be an instantaneous event affecting only one particle.

The collision–coupling idea is incorporated into the Hamiltonian model of dynamics (7.10) by adding a stochastic force \( R_i = R_i(t) \) to the \( \dot{p} \) equation

\[ \dot{q}^i = \partial_{p^i} H, \quad \dot{p}_i = -\partial_{q^i} H + R_i(t). \]

On the other hand, the so–called Ito stochastic integral represents a kind of classical Riemann–Stieltjes integral from linear functional analysis, which is (in 1D case) for an arbitrary time–function \( G(t) \) defined as the \textit{mean square limit}

\[ \int_{t_n}^t G(t) dW(t) = \lim_{n \to \infty} \left\{ \frac{1}{n} \sum_{i=1}^{n} G(t_{i-1})[W(t_i) - W(t_{i-1})] \right\}. \]

Now, the general ND Markov process can be defined by Ito stochastic differential equation (SDE),

\[ dx_i(t) = A_i[x^i(t), t]dt + B_{ij}[x^j(t), t] dW^j(t), \quad x^i(0) = x_0, \quad (i, j = 1, \ldots, N) \]
or corresponding *Ito stochastic integral equation*

\[ x^i(t) = x^i(0) + \int_0^t ds A_i[x^j(s), s] + \int_0^t dW^j(s) B_{ij}[x^j(s), s], \]

in which \( x^i(t) \) is the variable of interest, the vector \( A_i[x(t), t] \) denotes deterministic *drift*, the matrix \( B_{ij}[x(t), t] \) represents continuous stochastic *diffusion fluctuations*, and \( W^j(t) \) is an \( N \)-variable *Wiener process* (i.e., generalized Brownian motion) \([Wiener (1961)](WIENER1961)\), and \( dW^j(t) = W^j(t + dt) - W^j(t) \).

Now, there are three well–known special cases of the *Chapman–Kolmogorov equation* (see \([Gardiner (1985)](GARDINER1985)\)):

1. When both \( B_{ij}[x(t), t] \) and \( W(t) \) are zero, i.e., in the case of pure deterministic motion, it reduces to the *Liouville equation*

\[
\partial_t P(x', t'| x'', t'') = - \sum_i \frac{\partial}{\partial x^i} \{ A_i[x(t), t] P(x', t'| x'', t'') \} .
\]

2. When only \( W(t) \) is zero, it reduces to the *Fokker–Planck equation*

\[
\partial_t P(x', t'| x'', t'') = - \sum_i \frac{\partial}{\partial x^i} \{ A_i[x(t), t] P(x', t'| x'', t'') \} \\
+ \frac{1}{2} \sum_{ij} \frac{\partial^2}{\partial x^i \partial x^j} \{ B_{ij}[x(t), t] P(x', t'| x'', t'') \} .
\]

3. When both \( A_i[x(t), t] \) and \( B_{ij}[x(t), t] \) are zero, i.e., the state–space consists of integers only, it reduces to the *Master equation* of discontinuous jumps

\[
\partial_t P(x', t'| x'', t'') = \int dx \{ W(x'| x'', t) P(x', t'| x'', t'') - W(x'| x', t) P(x', t'| x'', t'') \} .
\]

The *Markov assumption* can now be formulated in terms of the conditional probabilities \( P(x^i, t_i) \): if the times \( t_i \) increase from right to left, the conditional probability is determined entirely by the knowledge of the most recent condition. Markov process is generated by a set of conditional probabilities whose probability–density \( P = P(x', t'| x'', t'') \) evolution obeys
the general Chapman–Kolmogorov integro–differential equation
\[ \partial_t P = - \sum_i \frac{\partial}{\partial x_i} \{ A_i [x(t), t] P \} + \frac{1}{2} \sum_{ij} \frac{\partial^2}{\partial x_i \partial x_j} \{ B_{ij} [x(t), t] P \} + \int dx \{ W(x'|x'', t) P - W(x''|x', t) P \} \]

including deterministic drift, diffusion fluctuations and discontinuous jumps (given respectively in the first, second and third terms on the r.h.s.).

It is this general Chapman–Kolmogorov integro–differential equation, with its conditional probability density evolution, \( P = P(x', t' | x'', t'') \), that we are going to model by various forms of the Feynman path integral, providing us with the physical insight behind the abstract (conditional) probability densities.

3.1.1.2 Quantum Probability Concept

An alternative concept of probability, the so–called quantum probability, is based on the following physical facts (elaborated in detail in this section):

1. The time–dependent Schrödinger equation represents a complex–valued generalization of the real–valued Fokker–Planck equation for describing the spatio–temporal probability density function for the system exhibiting continuous–time Markov stochastic process.
2. The Feynman path integral \( \int \Sigma \) is a generalization of the time–dependent Schrödinger equation, including both continuous–time and discrete–time Markov stochastic processes.
3. Both Schrödinger equation and path integral give ‘physical description’ of any system they are modelling in terms of its physical energy, instead of an abstract probabilistic description of the Fokker–Planck equation.

Therefore, the Feynman path integral \( \int \Sigma \), as a generalization of the time–dependent Schrödinger equation, gives a unique physical description for the general Markov stochastic process, in terms of the physically based generalized probability density functions, valid both for continuous–time and discrete–time Markov systems.

Basic consequence: a different way for calculating probabilities. The difference is rooted in the fact that sum of squares is different from the square of sums, as is explained in the following text.

Namely, in Dirac–Feynman quantum formalism, each possible route from the initial system state \( A \) to the final system state \( B \) is called a
history. This history comprises any kind of a route (see Figure 3.1), ranging from continuous and smooth deterministic (mechanical-like) paths to completely discontinues and random Markov chains (see, e.g., Gardiner (1985)). Each history (labelled by index $i$) is quantitatively described by a complex number $z_i$ called the ‘individual transition amplitude’. Its absolute square, $|z_i|^2$, is called the individual transition probability. Now, the total transition amplitude is the sum of all individual transition amplitudes, $\sum_i z_i$, called the sum–over–histories. The absolute square of this sum–over–histories, $|\sum_i z_i|^2$, is the total transition probability.

In this way, the overall probability of the system’s transition from some initial state $A$ to some final state $B$ is given not by adding up the probabilities for each history–route, but by ‘head–to–tail’ adding up the sequence of amplitudes making–up each route first (i.e., performing the sum–over–histories) – to get the total amplitude as a ‘resultant vector’, and then squaring the total amplitude to get the overall transition probability.

Fig. 3.1 Two ways of physical transition from an initial state $A$ to the corresponding final state $B$. (a) Classical physics proposes a single deterministic trajectory, minimizing the total system’s energy. (b) Quantum physics proposes a family of Markov stochastic histories, namely all possible routes from $A$ to $B$, both continuous–time and discrete–time Markov chains, each giving an equal contribution to the total transition probability.

1Recall that a complex number $z = x + iy$, where $i = \sqrt{-1}$ is the imaginary unit, $x$ is the real part and $y$ is the imaginary part, can be represented also in its polar form, $z = r(\cos \theta + i \sin \theta)$, where the radius vector in the complex–plane, $r = |z| = \sqrt{x^2 + y^2}$, is the modulus or amplitude, and angle $\theta$ is the phase; as well as in its exponential form $z = re^{i\theta}$. In this way, complex numbers actually represent 2D vectors with usual vector ‘head–to–tail’ addition rule.
3.1.1.3 Quantum Coherent States

Recall that a quantum coherent state is a specific kind of quantum state of the quantum harmonic oscillator whose dynamics most closely resemble the oscillating behavior of a classical harmonic oscillator. It was the first example of quantum dynamics when Erwin Schrödinger derived it in 1926 while searching for solutions of the Schrödinger equation that satisfy the correspondence principle. The quantum harmonic oscillator and hence, the coherent state, arise in the quantum theory of a wide range of physical systems. For instance, a coherent state describes the oscillating motion of the particle in a quadratic potential well. In the quantum electrodynamics and other bosonic quantum field theories they were introduced by the 2005 Nobel Prize winning work of R. Glauber in 1963 [Glauber (1963a); Glauber (1963b)]. Here the coherent state of a field describes an oscillating field, the closest quantum state to a classical sinusoidal wave such as a continuous laser wave.

In classical optics, light is thought of as electromagnetic waves radiating from a source. Specifically, coherent light is thought of as light that is emitted by many such sources that are in phase. For instance, a light bulb radiates light that is the result of waves being emitted at all the points along the filament. Such light is incoherent because the process is highly random in space and time. On the other hand, in a laser, light is emitted by a carefully controlled system in processes that are not random but interconnected by stimulation and the resulting light is highly ordered, or coherent. Therefore a coherent state corresponds closely to the quantum state of light emitted by an ideal laser. Semi-classically we describe such a state by an electric field oscillating as a stable wave. Contrary to the coherent state, which is the most wave-like quantum state, the Fock state (e.g., a single photon) is the most particle-like state. It is indivisible and contains only one quanta of energy. These two states are examples of the opposite extremes in the concept of wave-particle duality. A coherent state distributes its quantum-mechanical uncertainty equally, which means that the phase and amplitude uncertainty are approximately equal. Conversely, in a single-particle state the phase is completely uncertain.

Formally, the coherent state \( |\alpha\rangle \) is defined to be the eigenstate of the annihilation operator \( a \), i.e., \( a|\alpha\rangle = \alpha|\alpha\rangle \). Note that since \( a \) is not Hermitian, \( \alpha = |\alpha|e^{i\theta} \) is complex. \( |\alpha| \) and \( \theta \) are called the amplitude and phase of the state.

Physically, \( a|\alpha\rangle = \alpha|\alpha\rangle \) means that a coherent state is left unchanged
by the detection (or annihilation) of a particle. Consequently, in a coherent state, one has exactly the same probability to detect a second particle. Note, this condition is necessary for the coherent state’s Poisson detection statistics. Compare this to a single-particle’s Fock state: Once one particle is detected, we have zero probability of detecting another.

Now, recall that a Bose–Einstein condensate (BEC) is a collection of boson atoms that are all in the same quantum state. An approximate theoretical description of its properties can be derived by assuming the BEC is in a coherent state. However, unlike photons, atoms interact with each other so it now appears that it is more likely to be one of the squeezed coherent states (see Breitenbach et al. (1997)). In quantum field theory and string theory, a generalization of coherent states to the case of infinitely many degrees of freedom is used to define a vacuum state with a different vacuum expectation value from the original vacuum.

3.1.1.4 Dirac’s \( \langle \text{bra} | \text{ket} \rangle \) Transition Amplitude

Now, we are ready to move on into the realm of quantum mechanics. Recall that P. Dirac [Dirac (1949)] described behavior of quantum systems in terms of complex-valued ket–vectors \( |A \rangle \) living in the Hilbert space \( \mathcal{H} \), and their duals, bra–covectors \( \langle B | \) (i.e., 1–forms) living in the dual Hilbert space \( \mathcal{H}^* \). The Hermitian inner product of kets and bras, the bra–ket \( \langle B | A \rangle \), is a complex number, which is the evaluation of the ket \( |A \rangle \) by the bra \( \langle B | \). This complex number, say \( r e^{i\theta} \), represents the system’s transition amplitude from its initial state \( A \) to its final state \( B \) i.e.,

\[
\text{Transition Amplitude} = \langle B | A \rangle = r e^{i\theta}.
\]

That is, there is a process that can mediate a transition of a system from initial state \( A \) to the final state \( B \) and the amplitude for this transition equals \( \langle B | A \rangle = r e^{i\theta} \). The absolute square of the amplitude, \( |\langle B | A \rangle|^2 \) represents the transition probability. Therefore, the probability of a transition event equals the absolute square of a complex number, i.e.,

\[
\text{Transition Probability} = |\langle B | A \rangle|^2 = |r e^{i\theta}|^2.
\]

These complex amplitudes obey the usual laws of probability: when a \(^2\)

\(^2\)Transition amplitude is otherwise called probability amplitude, or just amplitude.

\(^3\)Recall that in quantum mechanics, complex numbers are regarded as the vacuum–state, or the ground–state, and the entire amplitude \( \langle b | a \rangle \) is a vacuum–to–vacuum amplitude for a process that includes the creation of the state \( a \), its transition to \( b \), and the annihilation of \( b \) to the vacuum once more.
Feynman Path Integrals

transition event can happen in alternative ways then we add the complex numbers,

\[ \langle B_1 | A_1 \rangle + \langle B_2 | A_2 \rangle = r_1 e^{i\theta_1} + r_2 e^{i\theta_2}, \]

and when it can happen only as a succession of intermediate steps then we multiply the complex numbers,

\[ \langle B | A \rangle = \langle B | c \rangle \langle c | A \rangle = (r_1 e^{i\theta_1})(r_2 e^{i\theta_2}) = r_1 r_2 e^{i(\theta_1 + \theta_2)}. \]

In general,

(1) The amplitude for \( n \) mutually alternative processes equals the sum \( \sum_{k=1}^{n} r_k e^{i\theta_k} \) of the amplitudes for the alternatives; and

(2) If transition from \( A \) to \( B \) occurs in a sequence of \( m \) steps, then the total transition amplitude equals the product \( \prod_{j=1}^{m} r_j e^{i\theta_j} \) of the amplitudes of the steps.

Formally, we have the so-called expansion principle, including both products and sums:

\[ \langle B | A \rangle = \sum_{i=1}^{n} \langle B | c_i \rangle \langle c_i | A \rangle. \quad (3.1) \]

3.1.1.5 Feynman’s Sum–over–Histories

Now, iterating the Dirac’s expansion principle (6.11) over a complete set of all possible states of the system, leads to the simplest form of the Feynman path integral, or, sum–over–histories. Imagine that the initial and final states, \( A \) and \( B \), are points on the vertical lines \( x = 0 \) and \( x = n + 1 \), respectively, in the \( x – y \) plane, and that \( (c(k)_{(k), k}) \) is a given point on the line \( x = k \) for \( 0 \leq k \leq m \) (see Figure 3.2). Suppose that the sum of projectors for each intermediate state is complete: Applying the completeness

\[ \sum_{i} |c_i\rangle \langle c_i| = 1 \]

In Dirac’s language, the completeness of intermediate states becomes the statement that a certain sum of projectors is equal to the identity. Namely, suppose that \( \sum |c_i\rangle \langle c_i| \) with \( \langle c_i | c_i \rangle = 1 \) for each \( i \). Then

\[ \langle b | a \rangle = \langle b | a \rangle = \langle b | \sum_{i} |c_i\rangle \langle c_i| |a\rangle = \sum_{i} \langle b | c_i\rangle \langle c_i | a \rangle. \]

We assume that following sum is equal to one, for each \( k \) from 1 to \( n - 1 \):

\[ |c(k)_{1}\rangle \langle c(k)_{1}| + \ldots + |c(k)_{m}\rangle \langle c(k)_{m}| = 1. \]
Fig. 3.2 Analysis of all possible routes from the source A to the detector B is simplified to include only double straight lines (in a plane).

Iteratively, we get the following expression for the transition amplitude:

\[ \langle B | A \rangle = \sum_{i_1} \ldots \sum_{i_n} \langle B | c_{i_1} \rangle \langle c_{i_1} | c_{i_2} \rangle \ldots \langle c_{i_n} | A \rangle, \]

where the sum is taken over all \( i(k) \) ranging between 1 and \( m \), and \( k \) ranging between 1 and \( n \). Each term in this sum can be construed as a combinatorial route from \( A \) to \( B \) in the 2D space of the \( x-y \) plane. Thus the transition amplitude for the system going from some initial state \( A \) to some final state \( B \) is seen as a summation of contributions from all the routes connecting \( A \) to \( B \).

Feynman used this description to produce his celebrated path integral expression for a transition amplitude (see, e.g., [Grosche and Steiner (1998); Schulman (1981)]). His path integral takes the form

\[ \text{Transition Amplitude} = \langle B | A \rangle = \int \mathcal{D}[x] e^{iS[x]}, \]

where the sum–integral \( \int \mathcal{D}[x] \) is taken over all possible routes \( x = x(t) \) from the initial point \( A = A(t_{\text{ini}}) \) to the final point \( B = B(t_{\text{fin}}) \), and \( S = S[x] \) is the classical action for a particle to travel from \( A \) to \( B \) along a given extremal path \( x \). In this way, Feynman took seriously Dirac’s conjecture interpreting the exponential of the classical action functional \( (\mathcal{D}e^{iS}) \), resembling a complex number \( (re^{i\theta}) \), as an elemental amplitude. By integrating this elementary amplitude, \( \mathcal{D}e^{iS} \), over the infinitude of all possible histories, we get the total system’s transition amplitude.\(^6\)

\(^6\)For the quantum physics associated with a classical (Newtonian) particle the action \( S \) is given by the integral along the given route from \( a \) to \( b \) of the difference \( T-V \) where \( T \) is the classical kinetic energy and \( V \) is the classical potential energy of the particle. The beauty of Feynman’s approach to quantum physics is that it shows the relationship between the classical and the quantum in a particularly transparent manner. Classical
3.1.1.6 The Basic Form of a Path Integral

In Feynman’s version of non-relativistic quantum mechanics, the time evolution \( \psi(x', t') \rightarrow \psi(x'', t'') \) of the wave function \( \psi = \psi(x, t) \) of the elementary 1D particle may be described by the integral equation [Grosche and Steiner (1998)]

\[
\psi(x'', t'') = \int_{\mathbb{R}} K(x'', x'; t'', t') \psi(x', t'),
\]

(3.3)

where the propagator or Feynman kernel \( K = K(x'', x'; t'', t') \) is defined through a limiting procedure,

\[
K(x'', x'; t'', t') = \lim_{\epsilon \to 0} A^{-N} \prod_{k=1}^{N-1} \int dx_k e^{i \sum_{j=0}^{N-1} \epsilon L(x_{j+1}, (x_{j+1} - x_j)/\epsilon)}.
\]

(3.4)

The time interval \( t'' - t' \) has been discretized into \( N \) steps of length \( \epsilon = (t'' - t')/N \), and the r.h.s. of (3.4) represents an integral over all piecewise motion corresponds to those regions where all nearby routes contribute constructively to the summation. This classical path occurs when the variation of the action is null. To ask for those paths where the variation of the action is zero is a problem in the calculus of variations, and it leads directly to Newton’s equations of motion (derived using the Euler–Lagrangian equations). Thus with the appropriate choice of action, classical and quantum points of view are unified.

Also, a discretization of the Schrodinger equation

\[ i\hbar \frac{d\psi}{dt} = -\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V\psi, \]

leads to a sum–over–histories that has a discrete path integral as its solution. Therefore, the transition amplitude is equivalent to the wave \( \psi \). The particle travelling on the \( x- \)axis is executing a one-step random walk, see Figure 3.3.
Fig. 3.4 A piecewise linear particle path contributing to the discrete Feynman propagator.

linear paths $x(t)$ of a ‘virtual’ particle propagating from $x'$ to $x''$, illustrated in Figure 3.4.

The prefactor $A^{-N}$ is a normalization and $L$ denotes the Lagrangian function of the particle. Knowing the propagator $G$ is tantamount to having solved the quantum dynamics. This is the simplest instance of a path integral, and is often written schematically as

$$K(x', t'; x'', t'') = \int_{\Sigma} D[x(t)] e^{iS[x(t)]},$$

where $D[x(t)]$ is a functional measure on the ‘space of all paths’, and the exponential weight depends on the classical action $S[x(t)]$ of a path. Recall also that this procedure can be defined in a mathematically clean way if we Wick–rotate the time variable $t$ to imaginary values $t \mapsto \tau = it$, thereby making all integrals real [Reed and Simon (1975)].

3.1.1.7 Adaptive Path Integral

Now, we can extend the Feynman sum–over–histories (3.2), by adding the synaptic–like weights $w^i = w^i(t)$ into the measure $D[x]$, to get the adaptive path integral:

$$\text{Adaptive Transition Amplitude} = \langle B|A \rangle_w = \oint D[w, x] e^{iS[x]},$$

(3.5)

where the adaptive measure $D[w, x]$ is defined by the weighted product (of discrete time steps)

$$D[w, x] = \lim_{n \to \infty} \prod_{t=1}^{n} w^i(t) \, dx^i(t).$$

(3.6)
In (3.6) the synaptic weights \( w^i = w^i(t) \) are updated by the unsupervised Hebbian–like learning rule [Hebb (1949)]:

\[
w^i(t + 1) = w^i(t) + \frac{\sigma}{\eta} (w^d - w^a(t)), \tag{3.7}
\]

where \( \sigma = \sigma(t), \eta = \eta(t) \) represent local signal and noise amplitudes, respectively, while superscripts \( d \) and \( a \) denote desired and achieved system states, respectively. Theoretically, equations (3.5–3.7) define an \( \infty \)-dimensional complex–valued neural network\(^7\). Practically, in a computer simulation we can use \( 10^7 \leq n \leq 10^8 \), approaching the number of neurons in the brain. Such equations are usually solved using Markov–Chain Monte–Carlo methods on parallel (cluster) computers (see, e.g., [Wehner and Wolfer (1983a); Wehner and Wolfer (1983b)]).

### 3.1.2 Path Integral History

#### 3.1.2.1 Extract from Feynman’s Nobel Lecture

In his Nobel Lecture, December 11, 1965, Richard (Dick) Feynman said that he and his PhD supervisor, John Wheeler, had found the action \( A = A[x; t_i, t_j] \), directly involving the motions of the charges only\(^5\):

\[
A[x; t_i, t_j] = m_i \int (\dot{x}_\mu(t_i))^2 dt_i + \frac{1}{2} \epsilon_{ij} \int \delta (I^2_{ij}) \dot{x}_\mu(t_i) \dot{x}_\mu(t_j) dt_i dt_j
\]

with \( i \neq j \) \( \tag{3.8} \)

\[
I^2_{ij} = \left[ x^i_\mu(t_i) - x^j_\mu(t_j) \right] \left[ x^i_\mu(t_i) - x^j_\mu(t_j) \right],
\]

where \( x^i_\mu = x^i_\mu(t_i) \) is the four–vector position of the \( i \)th particle as a function of the proper time \( t_i \), while \( \dot{x}^i_\mu(t_i) = dx^i_\mu(t_i)/dt_i \) is the velocity four–vector.

The first term in the action \( A[x; t_i, t_j] \) \( \tag{3.8} \) is the integral of the proper time \( t_i \), the ordinary action of relativistic mechanics of free particles of mass \( m_i \) (summation over \( \mu \)). The second term in the action \( A[x; t_i, t_j] \) \( \tag{3.8} \) represents the electrical interaction of the charges. It is summed over each pair of charges (the factor \( 1/2 \) is to count each pair once, the term \( i = j \) is omitted to avoid self–action). The interaction is a double integral over a delta function of the square of space–time interval \( I^2 \) between two points on

\(^5\)For details on complex–valued neural networks, see e.g., complex–domain extension of the standard backpropagation learning algorithm [Georgiou and Koutsougeras (1992); Benvenuto and Piazza (2002)].

\(^6\)Wheeler–Feynman Idea [Wheeler and Feynman (1949)] “The energy tensor can be regarded only as a provisional means of representing matter. In reality, matter consists of electrically charged particles.”
the paths. Thus, interaction occurs only when this interval vanishes, that is, along light cones (see [Wheeler and Feynman (1949)]).

Feynman comments here: “The fact that the interaction is exactly one-half advanced and half-retarded meant that we could write such a principle of least action, whereas interaction via retarded waves alone cannot be written in such a way. So, all of classical electrodynamics was contained in this very simple form.”

“...The problem is only to make a quantum theory, which has as its classical analog, this expression (3.8). Now, there is no unique way to make a quantum theory from classical mechanics, although all the textbooks make believe there is. What they would tell you to do, was find the momentum variables and replace them by \((\hbar/\text{i})\partial/\partial x\), but I couldn’t find a momentum variable, as there wasn’t any.”

“The character of quantum mechanics of the day was to write things in the famous Hamiltonian way (in the form of Schrödinger equation), which described how the wave function changes from instant to instant, and in terms of the Hamiltonian operator \(H\). If the classical physics could be reduced to a Hamiltonian form, everything was all right. Now, least action does not imply a Hamiltonian form if the action is a function of anything more than positions and velocities at the same moment. If the action is of the form of the integral of the Lagrangian \(L = L(\dot{x}, x)\), a function of the velocities and positions at the same time \(t\),

\[
S[x] = \int L(\dot{x}, x)\, dt, \tag{3.9}
\]

then you can start with the Lagrangian \(L\) and then create a Hamiltonian \(H\) and work out the quantum mechanics, more or less uniquely. But the action \(A[x; t_i, t_j]\) involves the key variables, positions (and velocities), at two different times \(t_i\) and \(t_j\) and therefore, it was not obvious what to do to make the quantum-mechanical analogue...”

So, Feynman was looking for the action integral in quantum mechanics. He says: “...I simply turned to Professor Jehle and said, ‘Listen, do you know any way of doing quantum mechanics, starting with action – where the action integral comes into the quantum mechanics?’ ‘No’, he said, ‘but Dirac has a paper in which the Lagrangian, at least, comes into quantum mechanics.’ What Dirac said was the following: There is in quantum mechanics a very important quantity which carries the wave function from one time to another, besides the differential equation but equivalent to it, a kind of a kernel, which we might call \(K(x', x)\), which carries the wave
function $\psi(x)$ known at time $t$, to the wave function $\psi(x')$ at time $t + \varepsilon$,

$$\psi(x', t + \varepsilon) = \int K(x', x) \psi(x, t) \, dx.$$ 

Dirac points out that this function $K$ was analogous to the quantity in classical mechanics that you would calculate if you took the exponential of $i\varepsilon$ multiplied by the Lagrangian $L(\dot{x}, x)$, imagining that these two positions $x, x'$ corresponded to $t$ and $t + \varepsilon$. In other words,

$$K(x', x) \text{ is analogous to } e^{i\varepsilon \frac{\dot{x}' - x}{\hbar}}.$$ 

So, Feynman continues: “What does he mean, they are analogous; what does that mean, analogous? What is the use of that?” Professor Jehle said, ‘You Americans! You always want to find a use for everything!” I said that I thought that Dirac must mean that they were equal. ‘No”, he explained, ‘he doesn’t mean they are equal.” “Well”, I said, ‘Let’s see what happens if we make them equal.”

“So, I simply put them equal, taking the simplest example where the Lagrangian is

$$L = \frac{1}{2} M \dot{x}^2 - V(x),$$ 

but soon found I had to put a constant of proportionality $N$ in, suitably adjusted. When I substituted for $K$ to get

$$\psi(x', t + \varepsilon) = \int N \exp \left[ i\frac{\varepsilon}{\hbar} L(\frac{x' - x}{\varepsilon}, x) \right] \psi(x, t) \, dx \quad (3.10)$$

and just calculated things out by Taylor series expansion, out came the Schrödinger equation. So, I turned to Professor Jehle, not really understanding, and said, ‘Well, you see, Dirac meant that they were proportional.” Professor Jehle’s eyes were bugging out – he had taken out a little notebook and was rapidly copying it down from the blackboard, and said, ‘No, no, this is an important discovery. You Americans are always trying to find out how something can be used. That’s a good way to discover things!” So, I thought I was finding out what Dirac meant, but, as a matter of fact, had made the discovery that what Dirac thought was analogous, was, in fact, equal. I had then, at least, the connection between the Lagrangian and quantum mechanics, but still with wave functions and infinitesimal times.”

“It must have been a day or so later when I was lying in bed thinking about these things, that I imagined what would happen if I wanted to
calculate the wave function at a finite interval later. I would put one of these factors $e^{i\varepsilon L}$ in here, and that would give me the wave functions the next moment, $t + \varepsilon$, and then I could substitute that back into (3.10) to get another factor of $e^{i\varepsilon L}$ and give me the wave function the next moment, $t + 2\varepsilon$, and so on and so on. In that way I found myself thinking of a large number of integrals, one after the other in sequence. In the integrand was the product of the exponentials, which was the exponential of the sum of terms like $\varepsilon L$. Now, $L$ is the Lagrangian and $\varepsilon$ is like the time interval $dt$, so that if you took a sum of such terms, that’s exactly like an integral. That’s like Riemann’s formula for the integral $\int Ldt$, you just take the value at each point and add them together. We are to take the limit as $\varepsilon \to 0$. Therefore, the connection between the wave function of one instant and the wave function of another instant a finite time later could be get by an infinite number of integrals (because $\varepsilon$ goes to zero), of exponential where $S$ is the action expression (3.9). At last, I had succeeded in representing quantum mechanics directly in terms of the action $S[x]$.

Fully satisfied, Feynman comments: “This led later on to the idea of the transition amplitude for a path: that for each possible way that the particle can go from one point to another in space–time, there’s an amplitude. That amplitude is $e$ to the power of $[i/\hbar$ times the action $S[x]$ for the path], i.e., $e^{iS[x]/\hbar}$. Amplitudes from various paths superpose by addition. This then is another, a third way, of describing quantum mechanics, which looks quite different from that of Schrödinger or Heisenberg, but which is equivalent to them.”

“...Now immediately after making a few checks on this thing, what we wanted to do, was to substitute the action $A[x; t_i, t_j]$ (3.8) for the other $S[x]$ (3.9). The first trouble was that I could not get the thing to work with the relativistic case of spin one–half. However, although I could deal with the matter only nonrelativistically, I could deal with the light or the photon interactions perfectly well by just putting the interaction terms of (3.8) into any action, replacing the mass terms by the non–relativistic $Ldt = 1/2M\dot{x}^2dt$,

$$A[x; t_i, t_j] = \frac{1}{2} \sum_i m_i \int (\dot{x}_i^2) dt_i + \frac{1}{2} \sum_{i\neq j} e_i e_j \int \int \delta(I_{ij}) \dot{x}_i^\mu(t_i) \dot{x}_j^\mu(t_j) dt_i dt_j.$$
cause very much trouble. It just meant developing a new idea. Instead of wave functions we could talk about this: that if a source of a certain kind emits a particle, and a detector is there to receive it, we can give the amplitude that the source will emit and the detector receive, $e^{iA[x; t_i, t_j]/\hbar}$.

We do this without specifying the exact instant that the source emits or the exact instant that any detector receives, without trying to specify the state of anything at any particular time in between, but by just finding the amplitude for the complete experiment. And, then we could discuss how that amplitude would change if you had a scattering sample in between, as you rotated and changed angles, and so on, without really having any wave functions... It was also possible to discover what the old concepts of energy and momentum would mean with this generalized action. And, so I believed that I had a quantum theory of classical electrodynamics – or rather of this new classical electrodynamics described by the action $A[x; t_i, t_j]$ (3.8) ...

3.1.2.2 Lagrangian Path Integral

Dirac and Feynman first developed the lagrangian approach to functional integration. To review this approach, we start with the time-dependent Schrödinger equation

$$i\hbar \partial_t \psi(x, t) = -\partial_x^2 \psi(x, t) + V(x) \psi(x, t)$$

appropriate to a particle of mass $m$ moving in a potential $V(x), x \in \mathbb{R}$. A solution to this equation can be written as an integral (see e.g., [Klauder (1997); Klauder (2000)]),

$$\psi(x'', t'') = \int K(x''', t'''; x', t') \psi(x', t') dx',$$

which represents the wave function $\psi(x'', t'')$ at time $t''$ as a linear superposition over the wave function $\psi(x', t')$ at the initial time $t', t' < t''$. The integral kernel $K(x''', t'''; x', t')$ is known as the propagator, and according to Feynman [Feynman (1948)] it may be given by

$$K(x''', t'''; x', t') = \mathcal{N} \int D[x] \, e^{i\phi/\hbar} \int [(m/2) \ddot{x}(t) - V(x(t))] dt,$$

which is a formal expression symbolizing an integral over a suitable set of paths. This integral is supposed to run over all continuous paths $x(t), t' \leq t \leq t''$, where $x(t'') = x''$ and $x(t') = x'$ are fixed end points for all paths. Note that the integrand involves the classical Lagrangian for the system.
To overcome the convergence problems, Feynman adopted a lattice regularization as a procedure to yield well-defined integrals which was then followed by a limit as the lattice spacing goes to zero called the continuum limit. With $\varepsilon > 0$ denoting the lattice spacing, the details regarding the lattice regularization procedure are given by

$$K(x'', t''; x', t') = \lim_{\varepsilon \to 0} \frac{(m/2\pi\hbar \varepsilon)^{(N+1)/2}}{N} \int \cdots \int \exp\left\{ \frac{i}{\hbar} \sum_{i=0}^{N} [(m/2\varepsilon)(x_{i+1} - x_i)^2 - \varepsilon V(x_i)] \right\} \prod_{i=1}^{N} dx_i,$$

where $x_{N+1} = x''$, $x_0 = x'$, and $\varepsilon \equiv (t'' - t')/(N + 1)$, $N \in \{1, 2, 3, \ldots\}$. In this version, at least, we have an expression that has a reasonable chance of being well defined, provided that one interprets the conditionally convergent integrals involved in an appropriate manner. One common and fully acceptable interpretation adds a convergence factor to the exponent of the preceding integral in the form $-(\varepsilon^2/2\hbar) \sum_{i=1}^{N} x_i^2$, which is a term that formally makes no contribution to the final result in the continuum limit save for ensuring that the integrals involved are now rendered absolutely convergent.

### 3.1.2.3 Hamiltonian Path Integral

It is necessary to retrace history at this point to recall the introduction of the phase–space path integral by [Feynman (1951); Grosche and Steiner (1998)]. In Appendix B to this article, Feynman introduced a formal expression for the configuration or $q$–space propagator given by (see e.g., [Klauder (1997); Klauder (2000)])

$$K(q'', t''; q', t') = \mathcal{M} \int \mathcal{D}[p] \mathcal{D}[q] \exp\left\{ \frac{i}{\hbar} \int [p \dot{q} - H(p, q)] dt \right\}.$$

In this equation one is instructed to integrate over all paths $q(t)$, $t' \leq t \leq t''$, with $q(t'') \equiv q''$ and $q(t') \equiv q'$ held fixed, as well as to integrate over all paths $p(t)$, $t' \leq t \leq t''$, without restriction.

It is widely appreciated that the phase–space path integral is more generally applicable than the original, Lagrangian, version of the path integral. For example, the original configuration space path integral is satisfactory for Lagrangians of the general form

$$L(x) = \frac{1}{2} m \ddot{x}^2 + A(x) \dot{x} - V(x),$$
but it is unsuitable, for example, for the case of a relativistic particle with
the Lagrangian

\[ L(x) = -m \sqrt{1 - \dot{x}^2} \]

expressed in units where the speed of light is unity. For such a system – as well as many more general expressions – the phase–space form of the path
integral is to be preferred. In particular, for the relativistic free particle, the phase–space path integral

\[ \mathcal{M} \int \mathcal{D}[p] \mathcal{D}[q] \exp \left\{ \frac{i}{\hbar} \int \left[ p \dot{q} - q \sqrt{1 - \dot{x}^2} + m^2 \right] dt \right\}, \]

is readily evaluated and induces the correct propagator.

### 3.1.2.4 Feynman–Kac Formula

Through his own research, M. Kac was fully aware of Wiener’s theory of Brownian motion and the associated diffusion equation that describes the corresponding distribution function. Therefore, it is not surprising that he was well prepared to give a path integral expression in the sense of Feynman for an equation similar to the time–dependent Schrödinger equation save for a rotation of the time variable by $-\pi/2$ in the complex–plane, namely, by the change $t \rightarrow -it$ (see e.g., [Klauder (1997); Klauder (2000)]. In particular, [Kac (1951)] considered the equation

\[ \partial_t \rho(x, t) = \partial_x^2 \rho(x, t) - V(x) \rho(x, t). \quad (3.11) \]

This equation is analogous to Schrödinger equation but differs from it in certain details. Besides certain constants which are different, and the change $t \rightarrow -it$, the nature of the dependent variable function $\rho(x, t)$ is quite different from the normal quantum mechanical wave function. For one thing, if the function $\rho$ is initially real it will remain real as time proceeds. Less obvious is the fact that if $\rho(x, t) \geq 0$ for all $x$ at some time $t$, then the function will continue to be nonnegative for all time $t$. Thus we can interpret $\rho(x, t)$ more like a probability density; in fact in the special case that $V(x) = 0$, then $\rho(x, t)$ is the probability density for a Brownian particle which underlies the Wiener measure. In this regard, $\nu$ is called the diffusion constant.

The fundamental solution of (3.11) with $V(x) = 0$ is readily given as

\[ W(x, T; y, 0) = \frac{1}{\sqrt{2\pi \nu T}} \exp \left( -\frac{(x - y)^2}{2\nu T} \right), \]
which describes the solution to the diffusion equation subject to the initial condition
\[ \lim_{T \to 0^+} W(x, T; y, 0) = \delta(x - y). \]

Moreover, it follows that the solution of the diffusion equation for a general initial condition is given by
\[ \rho(x'', t'') = \int W(x'', t''; x', t') \rho(x', t') \, dx'. \]

Iteration of this equation \( N \) times, with \( \epsilon = (t'' - t')/(N + 1) \), leads to the equation
\[ \rho(x'', t'') = N' \int \cdots \int e^{-(1/2\nu) \sum_{i=0}^{N} (x_{i+1} - x_i)^2} \prod_{i=1}^{N} dx_i \, \rho(x', t') \, dx', \]
where \( x_{N+1} \equiv x'' \) and \( x_0 \equiv x' \). This equation features the imaginary time propagator for a free particle of unit mass as given formally as
\[ W(x'', t''; x', t') = \mathcal{N} \int \mathcal{D}[x] \, e^{-(1/2\nu) \int \dot{x}^2 \, dt}, \]
where \( \mathcal{N} \) denotes a formal normalization factor.

The similarity of this expression with the Feynman path integral [for \( V(x) = 0 \)] is clear, but there is a profound difference between these equations. In the former (Feynman) case the underlying measure is only \textit{finitely additive}, while in the latter (Wiener) case the continuum limit actually defines a genuine measure, i.e., a \textit{countably additive measure} on paths, which is a version of the famous \textit{Wiener measure}. In particular,
\[ W(x'', t''; x', t') = \int d\mu^W_W(x), \]
where \( \mu^W_W \) denotes a measure on continuous paths \( x(t), t' \leq t \leq t'' \), for which \( x(t'') \equiv x'' \) and \( x(t') \equiv x' \). Such a measure is said to be a \textit{pinned} Wiener measure, since it specifies its path values at two time points, i.e., at \( t = t' \) and at \( t = t'' > t' \).

We note that Brownian motion paths have the property that with probability one they are concentrated on continuous paths. However, it is also true that the time derivative of a Brownian path is almost nowhere defined, which means that, with probability one, \( \dot{x}(t) = \pm \infty \) for all \( t \).
When the potential \( V(x) \neq 0 \) the propagator associated with (3.11) is formally given by

\[
W(x'', t''; x', t') = N \int D[x] e^{-(1/2\nu) \int \dot{x}^2 \, dt - \int V(x) \, dt},
\]

an expression which is well defined if \( V(x) \geq c, -\infty < c < \infty \). A mathematically improved expression makes use of the Wiener measure and reads

\[
W(x'', t''; x', t') = \int e^{-\int V(x(t)) \, dt} d\mu_V(x).
\]

This is an elegant relation in that it represents a solution to the differential equation (3.11) in the form of an integral over Brownian motion paths suitably weighted by the potential \( V \). Incidentally, since the propagator is evidently a strictly positive function, it follows that the solution of the differential equation (3.11) is nonnegative for all time \( t \) provided it is nonnegative for any particular time value.

3.1.2.5 \textit{Itô Formula}

\textit{Itô} (1960) proposed another version of a continuous–time regularization that resolved some of the troublesome issues. In essence, the proposal of \textit{Itô} takes the form given by

\[
\lim_{\nu \to \infty} N_{\nu} \int D[x] \exp\{(i/\hbar) \int [\frac{1}{2} \dot{x}^2 - V(x)] \, dt\} \exp\{- (1/2\nu) \int [\ddot{x}^2 + \dot{x}^2] \, dt\}.
\]

Note well the alternative form of the auxiliary factor introduced as a regulator. The additional term \( \ddot{x}^2 \), the square of the second derivative of \( x \), acts to smooth out the paths sufficiently well so that in the case of (21) both \( x(t) \) and \( \dot{x}(t) \) are continuous functions, leaving \( \ddot{x}(t) \) as the term which does not exist. However, since only \( x \) and \( \dot{x} \) appear in the rest of the integrand, the indicated path integral can be well defined; this is already a positive contribution all by itself (see e.g., \textit{Klauder} (1997), \textit{Klauder} (2000)).

3.1.3 \textit{Path–Integral Quantization}

3.1.3.1 \textit{Canonical versus Path–Integral Quantization}

Recall that in the usual, \textit{canonical formulation} of quantum mechanics, the system’s phase–space coordinates, \( q \), and momenta, \( p \), are replaced by the
corresponding Hermitian operators in the Hilbert space, with real measurable eigenvalues, which obey Heisenberg commutation relations.

The path–integral quantization is instead based directly on the notion of a propagator \( K(q_f, t_f; q_i, t_i) \) which is defined such that (see Ryder (1996) Cheng and Li (1984) Gunion (2003))

\[
\psi(q_f, t_f) = \int K(q_f, t_f; q_i, t_i) \psi(q_i, t_i) dq_i,
\]

i.e., the wave function \( \psi(q_f, t_f) \) at final time \( t_f \) is given by a Huygens principle in terms of the wave function \( \psi(q_i, t_i) \) at an initial time \( t_i \), where we have to integrate over all the points \( q_i \) since all can, in principle, send out little wavelets that would influence the value of the wave function at \( q_f \) at the later time \( t_f \). This equation is very general and is an expression of causality. We use the normal units with \( \hbar = 1 \).

According to the usual interpretation of quantum mechanics, \( \psi(q_f, t_f) \) is the probability amplitude that the particle is at the point \( q_f \) and the time \( t_f \), which means that \( K(q_f, t_f; q_i, t_i) \) is the probability amplitude for a transition from \( q_i \) and \( t_i \) to \( q_f \) and \( t_f \). The probability that the particle is observed at \( q_f \) at time \( t_f \) if it began at \( q_i \) at time \( t_i \) is

\[
P(q_f, t_f; q_i, t_i) = |K(q_f, t_f; q_i, t_i)|^2.
\]

Let us now divide the time interval between \( t_i \) and \( t_f \) into two, with \( t \) as the intermediate time, and \( q \) the intermediate point in space. Repeated application of (3.12) gives

\[
\psi(q_f, t_f) = \int \int K(q_f, t_f; q, t) dq K(q, t; q_i, t_i) \psi(q_i, t_i) dq_i,
\]

from which it follows that

\[
K(q_f, t_f; q_i, t_i) = \int dq K(q_f, t_f; q, t) K(q, t; q_i, t_i).
\]

This equation says that the transition from \( (q_i, t_i) \) to \( (q_f, t_f) \) may be regarded as the result of the transition from \( (q_i, t_i) \) to all available intermediate points \( q \) followed by a transition from \( (q, t) \) to \( (q_f, t_f) \). This notion of all possible paths is crucial in the path–integral formulation of quantum mechanics.

Now, recall that the state vector \( |\psi, t\rangle_S \) in the Schrödinger picture is related to that in the Heisenberg picture \( |\psi\rangle_H \) by

\[
|\psi, t\rangle_S = e^{-iHt} |\psi\rangle_H,
\]
or, equivalently,

$$|\psi\rangle_H = e^{iHt} |\psi, t\rangle_S.$$

We also define the vector

$$|q, t\rangle_H = e^{iHt} |q\rangle_S,$$

which is the Heisenberg version of the Schrödinger state $|q\rangle$. Then, we can equally well write

$$\psi(q, t) = \langle q, t | \psi \rangle_H. \tag{3.13}$$

By completeness of states we can now write

$$\langle q_f, t_f | \psi \rangle_H = \int \langle q_f, t_f | q_i, t_i \rangle_H \langle q_i, t_i | \psi \rangle_H dq_i,$$

which with the definition of (3.13) becomes

$$\psi(q_f, t_f) = \int \langle q_f, t_f | q_i, t_i \rangle_H \psi(q_i, t_i) dq_i. \tag{3.14}$$

Comparing with (3.12), we get

$$K(q_f, t_f; q_i, t_i) = \langle q_f, t_f | q_i, t_i \rangle_H.$$

Now, let us calculate the quantum–mechanics propagator

$$\langle q', t' | q, t \rangle_H = \langle q' | e^{-iH(t-t')} q \rangle,$$

using the path–integral formalism that will incorporate the direct quantization of the coordinates, without Hilbert space and Hermitian operators.

The first step is to divide up the time interval into $n + 1$ tiny pieces:

$$t_l = l\varepsilon + t \text{ with } t' = (n+1)\varepsilon + t.$$

Then, by completeness, we can write (dropping the Heisenberg picture index $H$ from now on)

$$\langle q', t' | q, t \rangle = \int dq_1(t_1) \ldots dq_n(t_n) \langle q', t' | q_n, t_n \rangle \times \langle q_n, t_n | q_{n-1}, t_{n-1} \rangle \ldots \langle q_1, t_1 | q, t \rangle. \tag{3.14}$$

The integral $\int dq_1(t_1) \ldots dq_n(t_n)$ is an integral over all possible paths, which are not trajectories in the normal sense, since there is no requirement of continuity, but rather Markov chains.

Now, for small $\varepsilon$ we can write

$$\langle q', \varepsilon | q, 0 \rangle = \langle q' | e^{-i\varepsilon H(P, Q)} | q \rangle = \delta(q' - q) - i\varepsilon \langle q' | H(P, Q) | q \rangle.$$
where $H(P,Q)$ is the Hamiltonian (e.g., $H(P,Q) = \frac{1}{2}P^2 + V(Q)$, where $P,Q$ are the momentum and coordinate operators). Then we have (see Ryder (1996), Cheng and Li (1984), Gunion (2003))

$$
\langle q' | H(P,Q) | q \rangle = \int \frac{dp}{2\pi} e^{ip(q' - q)} H \left( p, \frac{1}{2}(q' + q) \right).
$$

Putting this into our earlier form we get

$$
\langle q', \varepsilon | q, 0 \rangle \simeq \int \frac{dp}{2\pi} \exp \left\{ i \left( p(q' - q) - \varepsilon H \left( p, \frac{1}{2}(q' + q) \right) \right) \right\},
$$

where the 0th order in $\varepsilon \to 0$ and the 1st order in $\varepsilon \to -i\varepsilon \langle q' | H(P,Q) | q \rangle$. If we now substitute many such forms into (3.14) we finally get

$$
\langle q', t' | q, t \rangle \equiv \lim_{n \to \infty} \int \prod_{i=1}^{n} dq_i \prod_{k=1}^{n+1} \frac{dp_k}{2\pi} \quad (3.15)
$$

$$
\times \exp \left\{ i \sum_{j=1}^{n+1} \left[ p_j (q_j - q_{j-1}) \right] - H \left( p_j, \frac{1}{2}(q_j + q_{j+1}) \right) (t_j - t_{j-1}) \right\},
$$

with $q_0 = q$ and $q_{n+1} = q'$. Roughly, the above formula says to integrate over all possible momenta and coordinate values associated with a small interval, weighted by something that is going to turn into the exponential of the action $\exp S$ in the limit where $\varepsilon \to 0$. It should be stressed that the different $q_i$ and $p_k$ integrals are independent, which implies that $p_k$ for one interval can be completely different from the $p_{k'}$ for some other interval (including the neighboring intervals). In principle, the integral (3.15) should be defined by analytic continuation into the complex–plane of, for example, the $p_k$ integrals.

Now, if we go to the differential limit where we call $t_j - t_{j-1} \equiv d\tau$ and write \( \frac{(q_j - q_{j-1})}{(t_j - t_{j-1})} \equiv \dot{q} \), then the above formula takes the form

$$
\langle q', t' | q, t \rangle = \int \mathcal{D}[p] \mathcal{D}[q] \exp \left\{ \int_{t}^{t'} [p\dot{q} - H(p,q)] d\tau \right\},
$$

where we have used the shorthand notation

$$
\int \mathcal{D}[p] \mathcal{D}[q] \equiv \int \prod_{\tau} \frac{dq(\tau) dp(\tau)}{2\pi}.
$$
Note that the above integration is an integration over the \( p \) and \( q \) values at every time \( \tau \). This is what we call a functional integral. We can think of a given set of choices for all the \( p(\tau) \) and \( q(\tau) \) as defining a path in the 6D phase-space. The most important point of the above result is that we have get an expression for a quantum–mechanical transition amplitude in terms of an integral involving only pure complex numbers, without operators.

We can actually perform the above integral for Hamiltonians of the type \( H = H(P,Q) \). We use square completion in the exponential for this, defining the integral in the complex \( p \) plane and continuing to the physical situation. In particular, we have

\[
\int_{-\infty}^{\infty} dp \exp \left\{ \frac{i\varepsilon(p\dot{q} - \frac{1}{2}p^2)}{2\pi} \right\} = \frac{1}{\sqrt{2\pi i\varepsilon}} \exp \left[ \frac{1}{2}i\varepsilon\dot{q}^2 \right].
\]


\[
\langle q', t' | q, t \rangle = \lim_{n \to \infty} \int \prod_{i} dq_i \sqrt{2\pi i\varepsilon} \exp \left\{ i\varepsilon \sum_{j=1}^{n+1} \frac{1}{2} \left( \frac{q_j - q_{j-1}}{\varepsilon} \right)^2 - V(\frac{q_j + q_{j+1}}{2}) \right\}.
\]

This can be formally written as

\[
\langle q', t' | q, t \rangle = \int D[q] e^{iS[q]}, \quad \text{where}
\]

\[
\int D[q] \equiv \int \prod_{i} \frac{dq_i}{\sqrt{2\pi i\varepsilon}}, \quad \text{while} \quad S[q] = \int_{t}^{t'} L(q, \dot{q}) d\tau
\]

is the standard action with the Lagrangian

\[
L = \frac{1}{2} \dot{q}^2 - V(q).
\]

Generalization to many degrees of freedom is straightforward:

\[
\langle q_1'...q_N', t' | q_1...q_N, t \rangle = \int D[p]D[q] \exp \left\{ i \int_{t}^{t'} \left[ \sum_{n=1}^{N} p_n \dot{q}_n - H(p_n, q_n) \right] d\tau \right\}
\]

with \( \int D[p]D[q] = \int \prod_{n=1}^{N} dq_n dp_n \frac{2\pi}{2\pi} \).

Here, \( q_n(t) = q_n \) and \( q_n(t') = q'_n \) for all \( n = 1,...,N \), and we are allowing for the full Hamiltonian of the system to depend upon all the \( N \) momenta and coordinates collectively.
3.1.3.2 \textit{Elementary Applications}

(i) Consider first
\[\langle q', t' | Q(t_0) | q, t \rangle = \int \prod dq_i \langle q', t' | q_{i_0}, t_{i_0} \rangle \ldots \langle q_0, t_0 | Q(t_0) | q_{i-1}, t_{i-1} \rangle \ldots \langle q_1, t_1 | q, t \rangle,\]
where we choose one of the time interval ends to coincide with \(t_0\), i.e., \(t_{i_0} = t_0\). If we operate \(Q(t_0)\) to the left, then it is replaced by its eigenvalue \(q_{i_0} = q(t_0)\). Aside from this one addition, everything else is evaluated just as before and we will obviously get
\[\langle q', t' | Q(t_0) | q, t \rangle = \int D[p]D[q] q(t_0) \exp \left\{ i \int_{t_0}^{t'} [p \dot{q} - H(p, q)] d\tau \right\}.\]

(ii) Next, suppose we want a \textit{path–integral expression} for
\[\langle q', t' | Q(t_1) Q(t_2) | q, t \rangle\]
in the case where \(t_1 > t_2\). For this, we have to insert as intermediate states \(|q_{i_1}, t_{i_1}\rangle \langle q_{i_1}, t_{i_1}|\) with \(t_{i_1} = t_1\) and \(|q_{i_2}, t_{i_2}\rangle \langle q_{i_2}, t_{i_2}|\) with \(t_{i_2} = t_2\) and since we have ordered the times at which we do the insertions we must have the first insertion to the left of the 2nd insertion when \(t_1 > t_2\). Once these insertions are done, we evaluate \(\langle q_{i_1}, t_{i_1} | Q(t_1) = \langle q_{i_1}, t_{i_1}| q(t_1)\) and \(\langle q_{i_2}, t_{i_2} | Q(t_2) = \langle q_{i_2}, t_{i_2}| q(t_2)\) and then proceed as before and get
\[\langle q', t' | Q(t_1) Q(t_2) | q, t \rangle = \int D[p]D[q] q(t_1) q(t_2) \exp \left\{ i \int_{t}^{t'} [p \dot{q} - H(p, q)] d\tau \right\}.\]

Now, let us ask what the above integral is equal to if \(t_2 > t_1\)? It is obvious that what we get for the above integral is \(\langle q', t' | Q(t_2) Q(t_1) | q, t \rangle\). Clearly, this generalizes to an arbitrary number of \(Q\) operators.

(iii) When we enter into quantum field theory, the \(Q\)’s will be replaced by fields, since it is the fields that play the role of coordinates in the 2nd quantization conditions.

3.1.3.3 \textit{Sources}

The source is represented by modifying the Lagrangian:
\[L \rightarrow L + J(t)q(t).\]
Let us define $|0,t⟩^J$ as the ground state (vacuum) vector (in the moving frame, i.e., with the $e^{iHt}$ included) in the presence of the source. The required transition amplitude is

$$Z[J] \propto ⟨0, +\infty|0, -\infty⟩^J,$$

where the source $J = J(t)$ plays a role analogous to that of an electromagnetic current, which acts as a source of the electromagnetic field. In other words, we can think of the scalar product $J_\mu A^\mu$, where $J_\mu$ is the current from a scalar (or Dirac) field acting as a source of the potential $A^\mu$. In the same way, we can always define a current $J$ that acts as the source for some arbitrary field $φ$. $Z[J]$ (otherwise denoted by $W[J]$) is a functional of the current $J$, defined as (see [Ryder (1996); Cheng and Li (1984); Gunion (2003)])

$$Z[J] \propto \int D[p] D[q] \exp \left\{i \int_\tau^\tau [p(\tau) \dot{q}(\tau) - H(p,q) + J(\tau)q(\tau)]d\tau \right\},$$

with the normalization condition $Z[J = 0] = 1$. Here, the argument of the exponential depends upon the functions $q(\tau)$ and $p(\tau)$ and we then integrate over all possible forms of these two functions. So the exponential is a functional that maps a choice for these two functions into a number. For example, for a quadratically completable $H(p,q)$, the $p$ integral can be performed as a $q$ integral

$$Z[J] \propto \int D[q] \exp \left\{i \int_{-\infty}^{+\infty} \left(L + Jq + \frac{1}{2}i\varepsilon q^2\right) d\tau \right\},$$

where the addition to $H$ was chosen in the form of a convergence factor $-\frac{i}{2}\varepsilon q^2$.

3.1.3.4 Fields

Let us now treat the abstract scalar field $φ(x)$ as a coordinate in the sense that we imagine dividing space up into many little cubes and the average value of the field $φ(x)$ in that cube is treated as a coordinate for that little cube. Then, we go through the multi-coordinate analogue of the procedure we just considered above and take the continuum limit. The final result is

$$Z[J] \propto \int D[φ] \exp \left\{i \int d^4x \left(L(φ(x)) + J(x)φ(x) + \frac{1}{2}i\varepsilon φ^2\right) \right\},$$
where for $\mathcal{L}$ we would employ the *Klein–Gordon Lagrangian* form. In the above, the $dx_0$ integral is the same as $d\tau$, while the $d^3\vec{x}$ integral is summing over the sub–Lagrangians of all the different little cubes of space and then taking the continuum limit. $\mathcal{L}$ is the *Lagrangian density* describing the Lagrangian for each little cube after taking the many–cube limit (see [Ryder (1996)](#Ryder), [Cheng and Li (1984)](#ChengLi), [Gunion (2003)](#Gunion)) for the full derivation).

We can now introduce interactions, $\mathcal{L}_I$. Assuming the simple form of the Hamiltonian, we have

$$Z[J] \propto \int \mathcal{D}[\phi] \exp \left\{ i \int d^4x \left( \mathcal{L}(\phi(x)) + \mathcal{L}_I(\phi(x)) + J(x)\phi(x) \right) \right\},$$

again using the normalization factor required for $Z[J = 0] = 1$.

For example of Klein Gordon theory, we would use

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_I,$$

where $\partial_\mu \equiv \partial_{2\mu}$ and we can freely manipulate indices, as we are working in Euclidean space $\mathbb{R}^3$. In order to define the above $Z[J]$, we have to include a convergence factor $i\varepsilon \phi^2$,

$$\mathcal{L}_0 \to \frac{1}{2} [\partial_\mu \phi \partial^\mu \phi - \mu^2 \phi^2], \quad \mathcal{L}_I = \mathcal{L}_I(\phi),$$

so that

$$Z[J] \propto \int \mathcal{D}[\phi] \exp \{ i \int d^4x (\frac{1}{2} [\partial_\mu \phi \partial^\mu \phi - \mu^2 \phi^2 + i\varepsilon \phi^2] + \mathcal{L}_I(\phi(x)) + J(x)\phi(x)) \}$$

is the appropriate *generating function* in the free field theory case.

### 3.1.3.5 Gauges

In the path integral approach to quantization of the *gauge theory*, we implement *gauge fixing* by restricting in some manner or other the path integral over gauge fields $\int \mathcal{D}[A_\mu]$. In other words we will write instead

$$Z[J] \propto \int \mathcal{D}[A_\mu] \delta (\text{some gauge fixing condition}) \exp \left\{ i \int d^4x \mathcal{L}(A_\mu) \right\}.$$

A common approach would be to start with the *gauge condition*

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} (\partial^\mu A_\mu)^2,$$
where the electrodynamic field tensor is given by $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$, and calculate

$$Z[J] \propto \int \mathcal{D}[A_{\mu}] \exp \left\{ i \int d^4x \left[ \mathcal{L}(A_{\mu}(x)) + J_{\mu}(x)A^\mu(x) \right] \right\}$$

as the generating function for the vacuum expectation values of time ordered products of the $A_{\mu}$ fields. Note that $J_{\mu}$ should be conserved ($\partial^{\mu}J_{\mu} = 0$) in order for the full expression $\mathcal{L}(A_{\mu}) + J_{\mu}A^\mu$ to be gauge–invariant under the integral sign when $A_{\mu} \to A_{\mu} + \partial^{\mu}\Lambda$. For a proper approach, see [Ryder (1996); Cheng and Li (1984); Gunion (2003)].

3.1.3.6 Riemannian–Symplectic Geometries

In this subsection, following [Shabanov and Klauder (1998)], we describe path integral quantization on Riemannian–symplectic manifolds. Let $\hat{q}^j$ be a set of Cartesian coordinate canonical operators satisfying the Heisenberg commutation relations $[\hat{q}^j, \hat{q}^k] = i\omega^{jk}$. Here $\omega^{jk} = -\omega^{kj}$ is the canonical symplectic structure. We introduce the canonical coherent states as $|q\rangle \equiv e^{i\hat{q}^j\hat{q}^j} |0\rangle$, where $\omega^{jn}\omega^{nk} = \delta^k_j$, and $|0\rangle$ is the ground state of a harmonic oscillator with unit angular frequency. Any state $|\psi\rangle$ is given as a function on phase–space in this representation by $\langle q|\psi\rangle = \psi(q)$. A general operator $\hat{A}$ can be represented in the form $\hat{A} = \int dq a(q)|q\rangle\langle q|$, where $a(q)$ is the lower symbol of the operator and $dq$ is a properly normalized form of the Liouville measure. The function $A(q,q') = \langle q|\hat{A}|q'\rangle$ is the kernel of the operator.

The main object of the path integral formalism is the integral kernel of the evolution operator

$$K_t(q,q') = \langle q|e^{-i\hat{H}t}|q'\rangle = \int_{q(0)=q'}^{q(t)=q} \mathcal{D}[q] e^{\int_0^t dt \left( \frac{1}{2}q^j\omega^{jk}\dot{q}^k - h \right)}.$$  \hfill (3.16)

Here $\hat{H}$ is the Hamiltonian, and $h(q)$ its symbol. The measure formally implies a sum over all phase-space paths pinned at the initial and final points, and a Wiener measure regularization implies the following replacement

$$\mathcal{D}[q] \to \mathcal{D}[\mu_{\nu}(q)] = \mathcal{D}[q] e^{-\frac{1}{2} \int_0^t d\tau \dot{q}^2} = N_{\nu}(t) d\mu_{\nu}^t(q).$$  \hfill (3.17)

The factor $N_{\nu}(t)$ equals $2\pi e^{\nu t}/\nu$ for every degree of freedom, $d\mu_{\nu}^t(q)$ stands for the Wiener measure, and $\nu$ denotes the diffusion constant. We denote by $K^\nu_t(q,q')$ the integral kernel of the evolution operator for a finite $\nu$. 

The Wiener measure determines a stochastic process on the flat phase–
space. The integral of the symplectic 1–form \( \int \omega dq \) is a stochastic integral
that is interpreted in the Stratonovich sense. Under general coordinate
transformations \( q = q(\bar{q}) \), the Wiener measure describes the same stochastic
process on flat space in the curvilinear coordinates \( dq^2 = d\sigma(\bar{q})^2 \), so that the
value of the integral is not changed apart from a possible phase term. After
the calculation of the integral, the evolution operator kernel is get by taking
the limit \( \nu \to \infty \). The existence of this limit, and also the covariance under
general phase-space coordinate transformations, can be proved through the
operator formalism for the regularized kernel \( K_\nu t(q, q') \).

Note that the integral (3.16) with the Wiener measure inserted can be
regarded as an ordinary Lagrangian path integral with a complex action,
where the configuration space is the original phase–space and the Hamil-
tonian \( h(q) \) serves as a potential. Making use of this observation it is not
hard to derive the corresponding Schrödinger–like equation

\[
\frac{\partial}{\partial t} K_\nu t(q, q') = \left[ \frac{\nu}{2} \left( \partial_{q^j} + i \frac{1}{2} \omega_{jk} q^k \right)^2 - i h(q) \right] K_\nu t(q, q'),
\]

subject to the initial condition \( K_{\nu=0} t(q, q') = \delta(q-q') \), \( 0 < \nu < \infty \). One can
show that \( K_\nu t \to K_t \) as \( \nu \to \infty \) for all \( t > 0 \). The covariance under gen-
eral coordinate transformations follows from the covariance of the “kinetic”
energy of the Schrödinger operator in (3.18): The Laplace operator is re-
placed by the Laplace–Beltrami operator in the new curvilinear coordinates
\( q = q(\bar{q}) \), so the solution is not changed, but written in the new coordinates.
This is similar to the covariance of the ordinary Schrödinger equation and
the corresponding Lagrangian path integral relative to general coordinate
transformations on the configuration space: The kinetic energy operator
(the Laplace operator) in the ordinary Schrödinger equation gives a term
quadratic in time derivatives in the path integral measure which is sufficient
for the general coordinate covariance. We remark that the regularization
procedure based on the modified Schrödinger equation (3.18) applies to far
more general Hamiltonians than those quadratic in canonical momenta and
leading to the conventional Lagrangian path integral.

### 3.1.4 Statistical Mechanics via Path Integrals

The Feynman path integral turns out to provide an elegant way of doing
statistical mechanics, as the partition function can be written as a path–
integral.
Recall that the standard partition function is defined as
\[ Z = \sum_j e^{-\beta E_j}, \]  
where \( \beta = 1/k_B T \) and \( E_j \) is the energy of the state \( |j\rangle \). We can write
\[ Z = \sum_j \langle j | e^{-\beta H} | j \rangle = \text{Tr} e^{-\beta H}. \]

But recall the definition of the propagator [MacKenzie (2000)]
\[ K(q', T; q, 0) = \langle q' | e^{-iHT} | q \rangle. \]

Suppose we consider \( T \) to be a complex parameter, and consider it to be pure imaginary, so that we can write \( T = -i\beta \), where \( \beta \) is real. Then we have
\[ K(q', -i\beta; q, 0) = \langle q' | e^{-iH(-i\beta)} | q \rangle = \langle q'| e^{\beta H} \sum_j |j\rangle \langle j| q \rangle \]
\[ = \sum_j e^{\beta E_j} \langle q'| j \rangle \langle j | q \rangle = \sum_j e^{-\beta E_j} \langle j | q \rangle \langle q' | j \rangle. \]

Putting \( q' = q \) and integrating over \( q \), we get
\[ \int dq K(q, -i\beta; q, 0) = \sum_j e^{-\beta E_j} \langle j | \int dq |q\rangle \langle q | j \rangle = Z. \tag{3.20} \]

This is the central observation of this section: that the propagator evaluated at negative imaginary time is related to the partition function [MacKenzie (2000)].

We can easily work out an elementary example such as the harmonic oscillator. Recall the path integral for it,
\[ K(q', T; q, 0) = \left( \frac{m\omega}{2\pi i \sin \omega T} \right)^{1/2} \exp \left\{ \frac{i m\omega}{2\sin \omega T} \left( (q'^2 + q^2) \cos \omega T - 2q'q \right) \right\}. \]

We can put \( q' = q \) and \( T = -i\beta \):
\[ K(q, -i\beta; q, 0) = \left( \frac{m\omega}{2\pi \sinh(\beta\omega)} \right)^{1/2} \exp \left\{ - \frac{m\omega q^2}{\sinh(\beta\omega)} (\cosh(\beta\omega) - 1) \right\}. \]
The partition function is thus
\[
Z = \int dq K(q, -i\beta; q, 0) = \left(\frac{m\omega}{2\pi \sinh(\beta\omega)}\right)^{1/2}\sqrt{\frac{\pi}{\sinh(\beta\omega)}} \left(\frac{m\omega}{\sinh(\beta\omega)} - 1\right)^{-1/2}\left(\cosh(\beta\omega) - 1\right)
\]
\[
= \left[2(\cosh(\beta\omega) - 1)\right]^{-1/2} = \left[e^{\beta\omega/2} (1 - e^{-\beta\omega})\right]^{-1/2} = \left[2\left(e^{\beta\omega/2} - 1\right)\right]^{-1/2} = \left[2\left(e^{\beta\omega/2} - 1\right)\right]^{-1/2} = \left[\frac{2}{1 - e^{-\beta\omega}}\right]^{-1/2}.
\]
Putting \(\hbar\) back in, we get the standard result:
\[
Z = \sum_{j=0}^{\infty} e^{-\beta(j+1/2)\omega}\hbar\omega.
\]

We can rewrite the partition function in terms of a path–integral. In ordinary (real) time,
\[
K(q', T; q, 0) = \int Dq(t) \exp \left[i \int_0^T dt \left(\frac{m\dot{q}^2}{2} - V(q)\right)\right],
\]
where the integral is over all paths from \((q, 0)\) to \((q', T)\). With \(q' = q\),
\(T \to -i\beta\),
\[
K(q, -i\beta; q, 0) = \int Dq(t) \exp \left[i \int_0^{-i\beta} dt \left(\frac{m\dot{q}^2}{2} - V(q)\right)\right],
\]
where we now integrate along the negative imaginary time axis (Figure 3.5).

Let us now define a real variable for this integration, \(\tau = it\). \(\tau\) is called the imaginary time, since when the time \(t\) is imaginary, \(\tau\) is real. Then the integral over \(\tau\) is along its real axis: when \(t: 0 \to -i\beta\), then \(\tau: 0 \to \beta\). We can write \(q\) as a function of the variable \(\tau\): \(q(t) \to q(\tau)\); then \(\dot{q} = i dq/d\tau\). The propagator becomes
\[
K(q, -i\beta; q, 0) = \int Dq(\tau) \exp \int_0^{\beta} d\tau \left(\frac{m}{2} \left(\frac{dq}{d\tau}\right)^2 + V(q)\right). \tag{3.21}
\]
The integral is over all functions \(q(\tau)\) such that \(q(0) = q(\beta) = q\).

The result \eqref{3.21} is an ‘imaginary–time’ or Euclidean path integral, defined by associating to each path an amplitude (statistical weight) \(\exp -S_E\), where \(S_E\) is the so–called Euclidean action, obtained from the usual Minkowski action by changing the sign of the potential energy term.
3.1.5 Path–Integral Monte–Carlo Simulation

The Monte Carlo method came into being roughly around the same time as the Feynman path integral. Anecdotally, the idea of gaining insight into a complex phenomenon by making various trials and studying the proportions of the respective outcomes occurred to Stanislaw Ulam while playing solitaire during an illness in 1946. The immediate application was, the problem of neutron diffusion studied in Los Alamos at that time. The name of the procedure first appeared in print in a classic paper by Metropolis and Ulam in 1949 \cite{Metropolis1949}, where the authors explicitly mentioned that the method they presented as a statistical approach to the study of integro–differential equations would sometimes be referred to as the Monte Carlo method. In classical statistical mechanics it quickly became a standard calculational tool.

The object of interest in Monte Carlo evaluations of Feynman’s path integral is the quantum statistical partition function \( Z \), given, in operator language, as the trace of the density operator \( \exp(-\beta \hat{H}) \) of the canonical ensemble (\( \beta = 1/k_B T \)) associated with a Hamilton operator describing \( N \) particles of mass \( m_i \) moving under the influence of a potential \( V \),

\[
\hat{H} = \sum_{i=1}^{N} \frac{\hat{p}_i^2}{2m_i} + V(\hat{r}_1, \ldots, \hat{r}_N).
\]
Expressed as a Feynman integral, the density matrix elements read
\[
\langle r | \exp(-\beta \hat{H}) | r' \rangle = \int_{r(0)=r}^{r(\infty)=r'} D\tau \exp \left\{ -\frac{1}{\hbar} \int_0^\infty L \left( \{ \vec{r}_i(\tau), \vec{r}_i(\tau) \} \right) d\tau \right\} ,
\]
where \( r \equiv \{ \vec{r}_1, \ldots, \vec{r}_N \} \), \( L \) denotes the classical Lagrangian
\[
L \left( \{ \vec{r}_i(\tau), \vec{r}_i(\tau) \} \right) = \sum_{i=1}^N \frac{m_i}{2} \dot{r}_i^2 + V(\vec{r}_1, \ldots, \vec{r}_N(\tau))
\]
expressed in imaginary time \( \tau \). The particles are assumed to be distinguishable. To evaluate the trace, we only need to set \( r = r' \) and integrate over \( r \). To take into account Bose or Fermi statistics for indistinguishable particles, the partition function splits into a sum of the direct Boltzmann part and parts with permuted endpoints.

The right hand side of (3.22) is a path integral for the \( 3N \) functions \( r \). The idea of a Monte Carlo evaluation of this quantity is to sample these paths stochastically and to get (approximate) information about the quantum statistics of the system by averaging over the finite set of paths generated in the sampling process.

Monte Carlo data always come with error bars and, in general, the errors associated with numerical Monte Carlo data stem from two distinct sources. A systematic error of Monte Carlo evaluations of the path integral follows from the need to identify the paths by a finite amount of computer information. This can be done by discretizing the paths at some set of points in the interval \((0, \hbar \beta)\). For a single particle moving in one dimension, the simplest discrete time approximation for \( \epsilon = \hbar \beta / L \) reads
\[
\langle x | \exp(-\beta \hat{H}) | x' \rangle = \lim_{L \to \infty} \frac{1}{A} \prod_{j=1}^{L-1} \int \frac{dx_j}{A} \exp \left\{ -\frac{1}{\hbar} \sum_{j=1}^L \left[ \frac{m}{2} \left( x_j - x_{j-1} \right)^2 + \epsilon V(x_{j-1}) \right] \right\} ,
\]
where \( A = (2\pi \hbar \epsilon / m)^{1/2} \) and \( x_0 = x \) and \( x_L = x' \). Alternatively, one may expand the individual paths in terms of an orthogonal function basis, e.g.

\[\text{There have been attempts to apply the Monte Carlo method to path integrals also for real time. However, due to the oscillating exponential one then has to deal with problems of numerical cancellation, and it is much harder to obtain results of some numerical accuracy. Therefore, we will here restrict myself to Monte Carlo work in imaginary time.}\]
by the Fourier decomposition,

\[ x(\tau) = x + \frac{(x' - x)\tau}{\hbar\beta} + \sum_{k=1}^{\infty} a_k \sin \frac{k\pi\tau}{\hbar\beta}, \]

and express the density matrix as

\[
\langle x | \exp(-\beta \hat{H}) | x' \rangle = \lim_{L' \to \infty} J \exp \left\{ -\frac{m}{2\hbar^2\beta} (x - x')^2 \right\} \\
\times \int L' \prod_{k=1}^{L'} da_k \exp \left\{ -\frac{a_k^2}{2\sigma_k^2} \right\} \times \exp \left\{ -\frac{1}{\hbar} \int_0^{\beta} V(x(\tau)) d\tau \right\},
\]

where \( \sigma_k = \sqrt{\frac{2\hbar^2\beta}{m(\pi k)^2}} \) and \( J \) is the Jacobian of the transformation from the integral over all paths to the integral over all Fourier coefficients. A systematic error then arises from the loss of information by the finite number \( L \) of points \( x_i \) on the discretized time axis or by the finite number \( L' \) of Fourier components \( a_k \) that are taken into account in the Monte Carlo sampling of the paths.

The other error source of Monte Carlo data is the statistical error due to the finite number \( N_m \) of paths that form the sample used for evaluating the statistical averages. To make matters worse, the probability of configurations is, in general highly peaked, making an independent sampling of paths highly inefficient in most cases. The remedy is to introduce some way of ‘importance sampling’ where configurations are generated according to their probability given by the exponential in (3.22). Statistical averages may then be computed as simple arithmetic means. A way to achieve this is by constructing Markov chains where transition probabilities between configuration are constructed that allow to generate a new configuration from a given one such that in the limit of infinitely many configurations the correct probability distribution of paths results. A very simple and universally applicable algorithm to set up such a Markov chain is the Metropolis algorithm introduced in 1953 [Metropolis et al. (1953)].

Here a new configuration is obtained by looking at some configuration with only one variable changed and accepting or rejecting it for the sample on the basis of a simple rule that depends only on the respective energies of the two configurations. The advantages of importance sampling on the basis of Markov chains are obtained on the cost that, in general, successive configurations are not statistically independent but autocorrelated. The crucial quantity is the integrated autocorrelation time \( \tau_{\text{int}}^{\text{O}} \) of a quantity of inter-
\( \mathcal{O} = \langle \mathcal{O} \rangle \) with \( \mathcal{O} = (1/N_m) \sum_{i=1}^{N_m} \mathcal{O}_i \) and \( \mathcal{O}_i \) computed for each path \( i \) in the sample. It enters the statistical error estimate \( \Delta \mathcal{O} \) for expectation values of \( \mathcal{O} \) computed from a Monte Carlo sample of \( N_m \) autocorrelated configurations as

\[
\Delta \mathcal{O} = \sqrt{\frac{\sigma_{\mathcal{O}_i}^2}{N_m} 2^{\text{int}} \tau},
\]

where \( \sigma_{\mathcal{O}_i}^2 \) is the variance of \( \mathcal{O}_i \).

With Monte Carlo generated samples of Feynman paths one can thus ‘measure’ thermodynamic properties of quantum systems like the internal energy and the specific heat, but also gain more detailed information about correlation functions, probability distributions and the like. In the low–temperature limit, \( \beta \to \infty \), quantum mechanical ground state properties are recovered.

The feasibility of evaluating the quantum statistical partition function of many–particle systems by Monte Carlo sampling of paths was well established by the early eighties and the method began to be applied to concrete problems, in particular in the chemical physics literature. It had also become clear that the method had severe restrictions if numerical accuracy was called for. In addition to the statistical error inherent to the Monte Carlo method, a systematic error was unavoidably introduced by the necessary discretization of the paths. Attempts to improve the accuracy by algorithmic improvements to reduce both the systematic and the statistical errors were reported in subsequent years. The literature is abundant and rather than trying to review the field we will only indicate some pertinent paths of development.

In Fourier PIMC methods, introduced in 1983 in the chemical physics context by Doll and Freeman [Doll and Freeman (1984); Freeman and Doll (1984)], the systematic error arises from the fact that only a finite number of Fourier components are taken into account. Here the systematic error could be reduced by the method of partial averaging [Coalson et al. (1986)].

In discrete time approximations arising from the short–time propagator or, equivalently, the high–temperature Green’s function various attempts have been made to find more rapidly converging formulations. Among these are attempts to include higher terms in an expansion of the Wigner–Kirkwood type, i.e. an expansion in terms of \( \hbar^2 / 2m \). Taking into account the first term of such an expansion would imply to replace the potential term \( \epsilon V(x_{j-1}) \) in (3.23) by [Raedt and Raedt (1985); Li and Broughton (1987)]
Kono et al. (1988)

\[ \epsilon V(x_{j-1}) \rightarrow \frac{\epsilon}{x-x'} \int_x^{x'} V(y) \, dy. \]

This improves the convergence of the density matrix (3.23) (from even less than \( O(1/L) \)) to \( O(1/L^2) \). For the full partition function, the convergence of the simple discretization scheme is already of order \( O(1/L^2) \) since due to the cyclic property of the trace, the discretization \( \epsilon V(x_{j-1}) \) is then equivalent to a symmetrized potential term \( \epsilon (V(x_{j-1}) + V(x_j))/2 \).

The convergence behavior of these formulations follows from the Trotter decomposition formula

\[ e^{-(A+B)} = \left[ e^{-\frac{\pi}{L}} e^{-\frac{\pi}{L}} \right]^L + O \left( \frac{1}{L} \right) = \left[ e^{-\frac{\pi}{L}} e^{-\frac{\pi}{L}} e^{-\frac{\pi}{L}} \right]^L + O \left( \frac{1}{L^2} \right), \]

valid for non-commuting operators \( A \) and \( B \) in a Banach space, identifying \( A \) with the kinetic energy \( \beta \sum \hat{p}_i^2/2m_i \) and \( B \) with the potential energy \( \beta V(\{\hat{x}_i\}) \). More rapidly converging discretization schemes were investigated on the basis of higher-order decompositions. Unfortunately, a direct, ‘fractal’ decomposition Suzuki (1990) of the form

\[ e^{-(A+B)} = \lim_{L \to \infty} \left[ e^{\alpha_1 \frac{\pi}{L}} e^{\beta_1 \frac{\pi}{L}} e^{\alpha_2 \frac{\pi}{L}} e^{\beta_2 \frac{\pi}{L}} \ldots \right]^L, \quad \sum \alpha_i = \sum \beta_i = 1, \]

inevitably leads to negative coefficients for higher decompositions and is therefore not amenable to Monte Carlo sampling of paths Janke and Sauer (1992). Higher–order Trotter decomposition schemes involving commutators have proven to be more successful Raedt and Raedt (1983). In particular, a decomposition of the form

\[ Z = \lim_{L \to \infty} \text{Tr} \left[ e^{-\frac{\pi}{L}} e^{-\frac{\pi}{L}} e^{-\frac{\pi}{2mL^2}} e^{-\frac{\pi}{2mL^2}} e^{-\frac{\pi}{L}} e^{-\frac{\pi}{L}} \right]^L, \]

derivable by making use of the cyclic property of the trace, is convergent of order \( O(1/L^4) \) and amounts to simply replacing the potential \( \epsilon V \) in (3.23) by an effective potential

\[ V_{\text{eff}} = V + \frac{(\beta \hbar)^2}{24mL^2} (V')^2. \]

Another problem for the numerical accuracy of PIMC simulations arises from the analog of the critical slowing down problem well–known for local update algorithms at second–order phase transitions in the simulation of
spin systems and lattice field theory. Since the correlations $\langle x_j x_{j+k} \rangle$ between variables $x_j$ and $x_{j+k}$ in the discrete time approximation only depend on the temperature and on the gaps between the energy levels and not, or at least not appreciably, on the discretization parameter $\epsilon$, the correlation length $\zeta$ along the discretized time axis always diverges linearly with $L$ when measured in units of the lattice spacing $\epsilon$. Hence in the continuum limit of $\epsilon \to 0$ with $\beta$ fixed or, equivalently, of $L \to \infty$ for local, importance sampling update algorithms, like the standard Metropolis algorithm, a slowing down occurs because paths generated in the Monte Carlo process become highly correlated. Since for simulations using the Metropolis algorithm autocorrelation times diverge as $\tau_{\text{int}} \propto L^z$ with $z \approx 2$ the computational effort (CPU time) to achieve comparable numerical accuracy in the continuum limit $L \to \infty$ diverges as $L \times L^z = L^{z+1}$.

To overcome this drawback, ad hoc algorithmic modifications like introducing collective moves of the path as a whole between local Metropolis updates were introduced then and again. One of the earliest more systematic and successful attempts to reduce autocorrelations between successive path configurations was introduced by Pollock and Ceperley (1984). Rewriting the discretized path integral, their method essentially amounts to a recursive transformation of the variables $x_i$ in such a way that the kinetic part of the energy can be taken care of by sampling direct Gaussian random variables and a Metropolis choice is made for the potential part. The recursive transformation can be done between some fixed points of the discretized paths, and the method has been applied in such a way that successively finer discretizations of the path were introduced between neighboring points. Invoking the polymer analog of the discretized path this method was christened the staging algorithm by Sprik et al. (1985).

The staging algorithm decorrelates successive paths very effectively because the whole staging section of the path is essentially sampled independently. In 1993, another explicitly non–local update was applied to PIMC simulations Janke and Sauer (1993) by transferring the so–called multigrid method known from the simulation of spin systems. Originating in the theory of numerical solutions of partial differential equations, the idea of the multigrid method is to introduce a hierarchy of successively coarser grids in order to take into account long wavelength fluctuations more effectively. Moving variables of the coarser grids then amounts to a collective move of neighbouring variables of the finer grids, and the formulation allows to give a recursive description of how to cycle most effectively through the various levels of the multigrid. Particularly successful is the so–called W–cycle.
Both the staging algorithm and the multigrid W–cycle have been shown to beat the slowing down problem in the continuum limit completely by reducing the exponent $z$ to $z \approx 0$ [Janke and Sauer (1996)].

Another cause of severe correlations between paths arises if the probability density of configurations is sharply peaked with high maxima separated by regions of very low probability density. In the statistical mechanics of spin systems this is the case at a first-order phase transition. In PIMC simulations the problem arises for tunneling situations like, e.g., for a double well potential with a high potential barrier between the two wells. In these cases, an unbiased probing of the configuration space becomes difficult because the system tends to get stuck around one of the probability maxima. A remedy to this problem is to simulate an auxiliary distribution that is flat between the maxima and to recover the correct Boltzmann distribution by an appropriate reweighting of the sample. The procedure is known under the name of "umbrella sampling" or "multicanonical sampling". It was shown to reduce autocorrelations for PIMC simulations of a single particle in a 1D double well, and it can also be combined with multigrid acceleration [Janke and Sauer (1994)].

The statistical error associated with a Monte Carlo estimate of an observable $O$ cannot only be reduced by reducing autocorrelation times $\tau_{int}^O$. If the observable can be measured with two different estimators $U_i$ that yield the same mean $U_i^{(L)} = \langle U_i \rangle$ with $O = \lim_{L \to \infty} U_i^{(L)}$, the estimator with the smaller variance $\sigma_i^2$ is to be preferred. Straightforward differentiation of the discretized path integral (3.23) leads to an estimator of the energy that explicitly measures the kinetic and potential parts of the energy by

$$U_k = \frac{L}{2\beta} - \frac{m}{2L} \sum_j \left( \frac{x_j - x_{j-1}}{\epsilon} \right)^2 + \frac{1}{L} \sum_{i=1}^L V(x_i).$$

The variance of this so–called kinetic–energy estimator diverges with $L$. Another estimator can be derived by invoking the path analog of the virial theorem

$$\frac{L}{2\beta} - \frac{m}{2} \left\langle \left( \frac{x_j - x_{j-1}}{\epsilon} \right)^2 \right\rangle = \frac{1}{2} \langle x_j V'(x_j) \rangle,$$

and the variance of the virial estimator

$$U_v = \frac{1}{2L} \sum_{i=1}^L x_i V'(x_i) + \frac{1}{L} \sum_{i=1}^L V(x_i).$$
Quantum Leap

does not depend on $L$. In the early eighties, investigations of the ‘kinetic’ and the ‘virial’ estimators focussed on their variances [Parrinello and Rahman (1984)]. Some years later, it was pointed out that a correct assessment of the accuracy also has to take into account the autocorrelations, and it was demonstrated that for a standard Metropolis simulation of the harmonic oscillator the allegedly less successful ‘kinetic’ estimator gave smaller errors than the ‘virial’ estimator. In 1989 it was shown [Cao and Berne (1989)] that conclusions about the accuracy also depend on the particular Monte Carlo update algorithm at hand since modifications of the update scheme such as inclusion of collective moves of the whole path affect the autocorrelations of the two estimators in a different way. A careful comparison of the two estimators which disentangles the various factors involved was given in [Janke and Sauer (1997)]. Here it was also shown that a further reduction of the error may be achieved by a proper combination of both estimators without extra cost.

Application of the Monte Carlo method to quantum systems is not restricted to direct sampling of Feynman paths but this method has attractive features. It is not only conceptually suggestive but also allows for algorithmic improvements that help to make the method useful even when the problems at hand require considerable numerical accuracy. However, algorithmic improvements like the ones alluded to above have tended to be proposed and tested mainly for simple, one–particle systems. On the other hand, the power of the Monte Carlo method is, of course, most welcome in those cases where analytical methods fail. For more complicated systems, however, evaluation of the algorithms and control of numerical accuracy is also more difficult.

For more details on path–integral Monte Carlo techniques, see [Sauer (2001)].

3.1.6 Sum over Geometries and Topologies

Recall that the term quantum gravity (or quantum geometrodynamics, or quantum geometry), is usually understood as a consistent fundamental quantum description of gravitational space–time geometry whose classical limit is Einstein’s general relativity. Among the possible ramifications of such a theory are a model for the structure of space–time near the Planck scale, a consistent calculational scheme to calculate gravitational effects at all energies, a description of quantum geometry near space–time singularities and a non–perturbative quantum description of
4D black holes. It might also help us in understanding cosmological issues about the beginning and end of the universe, i.e., the so-called ‘big bang’ and ‘Big–Crunch’ (see e.g., [Penrose (1967); Penrose (1994); Penrose (1997)].

From what we know about the quantum dynamics of other fundamental interactions it seems eminently plausible that also the gravitational excitations should at very short scales be governed by quantum laws. Now, conventional perturbative path integral expansions of gravity, as well as perturbative expansion in the string coupling in the case of unified approaches, both have difficulty in finding any direct or indirect evidence for quantum gravitational effects, be they experimental or observational, which could give a feedback for model building. The outstanding problems mentioned above require a non–perturbative treatment; it is not sufficient to know the first few terms of a perturbation series. The real goal is to search for a non–perturbative definition of such a theory, where the initial input of any fixed ‘background metric’ is inessential (or even undesirable), and where ‘space–time’ is determined dynamically. Whether or not such an approach necessarily requires the inclusion of higher dimensions and fundamental supersymmetry is currently unknown (see [Ambjørn and Kristjansen (1993); Ambjørn and Loll (1998); Ambjørn et al. (2000a); Ambjørn et al. (2000b); Ambjørn et al. (2001a); Ambjørn et al. (2001b); Ambjørn et al. (2001d); Dasgupta and Loll (2001)]).

Such a non–perturbative viewpoint is very much in line with how one proceeds in classical geometrodynamics, where a metric space–time \((M, g_{\mu\nu})\) (+ matter) emerges only as a solution to the familiar Einstein equation

\[
G_{\mu\nu}[g] \equiv R_{\mu\nu}[g] - \frac{1}{2} g_{\mu\nu} R[g] = -8\pi T_{\mu\nu} \Phi, \tag{3.24}
\]

which define the classical dynamics of fields \(\Phi = \Phi^{\mu\nu}\) on the space \(\mathcal{M}(M)\), the space of all metrics \(g = g_{\mu\nu}\) on a given smooth manifold \(M\). The analogous question we want to address in the quantum theory is: Can we get ‘quantum space–time’ as a solution to a set of non–perturbative quantum equations of motion on a suitable quantum analogue of \(\mathcal{M}(M)\) or rather, of the space of geometries, Geom\((M) = \mathcal{M}(M)/\text{Diff}(M)\)?

Now, this is not a completely straightforward task. Whichever way we want to proceed non–perturbatively, if we give up the privileged role of a flat, Minkowskian background space–time on which the quantization is to take place, we also have to abandon the central role usually played by the
Poincaré group, and with it most standard quantum field-theoretic tools for regularization and renormalization. If one works in a continuum metric formulation of gravity, the symmetry group of the Einstein–Hilbert action is instead the group $\text{Diff}(M)$ of diffeomorphisms on $M$, which in terms of local charts are the smooth invertible coordinate transformations $x^\mu \mapsto y^\mu(x^\nu)$.

In the following, we will describe a non-perturbative path integral approach to quantum gravity, defined on the space of all geometries, without distinguishing any background metric structure [Loll (2001)]. This is closely related in spirit with the canonical approach of loop quantum gravity [Rovelli (1998)] and its more recent incarnations using so-called spin networks (see, e.g., [Oriti (2001)]). ‘Non-perturbative’ here means in a covariant context that the path sum or integral will have to be performed explicitly, and not just evaluated around its stationary points, which can only be achieved in an appropriate regularization. The method we will employ uses a discrete lattice regularization as an intermediate step in the construction of the quantum theory.

3.1.6.1 Simplicial Quantum Geometry

In this section we will explain how one may construct a theory of quantum gravity from a non-perturbative path integral, using the method of Lorentzian dynamical triangulations. The method is minimal in the sense of employing standard tools from quantum field theory and the theory of critical phenomena and adapting them to the case of generally covariant systems, without invoking any symmetries beyond those of the classical theory. At an intermediate stage of the construction, we use a regularization in terms of simplicial Regge geometries, that is, piecewise linear manifolds. In this approach, ‘computing the path integral’ amounts to a conceptually simple and geometrically transparent ‘counting of geometries’, with additional weight factors which are determined by the EH action. This is done first of all at a regularized level. Subsequently, one searches for interesting continuum limits of these discrete models which are possible candidates for theories of quantum gravity, a step that will always involve a renormalization. From the point of view of statistical mechanics, one may think of Lorentzian dynamical triangulations as a new class of statistical models of Lorentzian random surfaces in various dimensions, whose building blocks are flat simplices which carry a ‘time arrow’, and whose dynamics is entirely governed by their intrinsic geometric properties.

Before describing the details of the construction, it may be helpful to
recall the path integral representation for a 1D non–relativistic particle (see previous subsection). The time evolution of the particle’s wave function $\psi$ may be described by the integral equation (3.3) above, where the propagator, or the Feynman kernel $G$, is defined through a limiting procedure (3.4). The time interval $t'' - t'$ has been discretized into $N$ steps of length $\epsilon = (t'' - t')/N$, and the r.h.s. of (3.4) represents an integral over all piecewise linear paths $x(t)$ of a ‘virtual’ particle propagating from $x'$ to $x''$, illustrated in Figure 3.4 above.

The prefactor $A^{-N}$ is a normalization and $L$ denotes the Lagrangian function of the particle. Knowing the propagator $G$ is tantamount to having solved the quantum dynamics. This is the simplest instance of a path integral, and is often written schematically as

$$G(x', t'; x'', t'') = \oint \mathcal{D}[x(t)] e^{iS[x(t)]},$$

(3.25)

where $\mathcal{D}[x(t)]$ is a functional measure on the ‘space of all paths’, and the exponential weight depends on the classical action $S[x(t)]$ of a path. Recall also that this procedure can be defined in a mathematically clean way if we Wick–rotate the time variable $t$ to imaginary values $t \mapsto \tau = i t$, thereby making all integrals real [Reed and Simon (1975)].

Can a similar strategy work for the case of Einstein geometrodynamics? As an analogue of the particle’s position we can take the geometry $[g_{ij}(x)]$ (i.e., an equivalence class of spatial metrics) of a constant–time slice. Can one then define a gravitational propagator

$$G([g'_{ij}], [g''_{ij}]) = \oint_{\text{Geom}(M)} \mathcal{D}[g_{\mu\nu}] e^{iS_{\text{EH}}[g_{\mu\nu}]},$$

(3.26)

from an initial geometry $[g']$ to a final geometry $[g'']$ (Figure 3.6) as a limit of some discrete construction analogous to that of the non-relativistic particle (3.4)? And crucially, what would be a suitable class of ‘paths’, that is, space–times $[g_{\mu\nu}]$ to sum over?

Now, to be able to perform the integration $\oint \mathcal{D}[g_{\mu\nu}]$ in a meaningful way, the strategy we will be following starts from a regularized version of the space $\text{Geom}(M)$ of all geometries. A regularized path integral $G(a)$ can be defined which depends on an ultraviolet cutoff $a$ and is convergent in a non–trivial region of the space of coupling constants. Taking the continuum limit corresponds to letting $a \to 0$. The resulting continuum theory – if it can be shown to exist – is then investigated with regard to its geometric properties and in particular its semiclassical limit.
3.1.6.2 Discrete Gravitational Path Integrals

Trying to construct non-perturbative path integrals for gravity from sums over discretized geometries, using approach of Lorentzian dynamical triangulations, is not a new idea. Inspired by the successes of lattice gauge theory, attempts to describe quantum gravity by similar methods have been popular on and off since the late 70’s. Initially the emphasis was on gauge-theoretic, first-order formulations of gravity, usually based on (compactified versions of) the Lorentz group, followed in the 80’s by ‘quantum Regge calculus’, an attempt to represent the gravitational path integral as an integral over certain piecewise linear geometries (see [Williams (1997)] and references therein), which had first made an appearance in approximate descriptions of classical solutions of the Einstein equations. A variant of this approach by the name of ‘dynamical triangulation(s)’ attracted a lot of interest during the 90’s, partly because it had proved a powerful tool in describing 2D quantum gravity (see the textbook [Ambjørn et al. (1997)] and lecture notes [Ambjørn et al. (2000a)] for more details).

The problem is that none of these attempts have so far come up with convincing evidence for the existence of an underlying continuum theory of 4D quantum gravity. This conclusion is drawn largely on the basis of numerical simulations, so it is by no means water-tight, although one can make an argument that the ‘symptoms’ of failure are related in the various approaches [Loll (1998)]. What goes wrong generically seems to be a dominance in the continuum limit of highly degenerate geometries, whose precise form depends on the approach chosen. One would expect that non-smooth geometries play a decisive role, in the same way as it can be shown in the particle case that the support of the measure in the continuum limit
is on a set of nowhere differentiable paths. However, what seems to happen in the case of the path integral for 4–geometries is that the structures get are too wild, in the sense of not generating, even at coarse–grained scales, an effective geometry whose dimension is anywhere near four.

The schematic phase diagram of Euclidean dynamical triangulations shown in Figure 3.7 gives an example of what can happen. The picture turns out to be essentially the same in both three and four dimensions: the model possesses infinite-volume limits everywhere along the critical line $k_0^{\text{crit}}(k_0)$, which fixes the bare cosmological constant as a function of the inverse Newton constant $k_0 \sim G_N^{-1}$. Along this line, there is a critical point $k_0^{\text{crit}}$ (which we now know to be of first–order in $d = 3, 4$) below which geometries generically have a very large effective or Hausdorff dimension. Above $k_0^{\text{crit}}$ we find the opposite phenomenon of ‘polymerization’: a typical element contributing to the state sum is a thin branched polymer, with one or more dimensions ‘curled up’ such that its effective dimension is around two.

![Fig. 3.7](image.png)

\textbf{Fig. 3.7} The phase diagram of 3D and 4D Euclidean dynamical triangulations (see text for explanation).

This problem has to do with the fact that the gravitational action is unbounded below, causing potential havoc in Euclidean versions of the path integral. Namely, what all the above-mentioned approaches have in common is that they work from the outset with Euclidean geometries, and associated Boltzmann-type weights $\exp(-S_{\text{eu}})$ in the path integral. In other words, they integrate over ‘space–times’ which know nothing about time, light cones and causality. This is done mainly for technical reasons, since

---

10In terms of geometry, this means that there are a few vertices at which the entire space–time ‘condenses’ in the sense that almost every other vertex in the simplicial space–time is about one link-distance away from them.
it is difficult to set up simulations with complex weights and since until recently a suitable Wick rotation was not known.

‘Lorentzian dynamical triangulations’, first proposed in [Ambjørn and Loll (1998)] and further elaborated in [Ambjørn et al. (2000b); Ambjørn et al. (2001a)], tries to establish a logical connection between the fact that non-perturbative path integrals were constructed for Euclidean instead of Lorentzian geometries and their apparent failure to lead to an interesting continuum theory.

3.1.6.3 Regge Calculus

The use of simplicial methods in general relativity goes back to the pioneering work of Regge [Regge (1961)]. In classical applications one tries to approximate a classical space–time geometry by a triangulation, that is, a piecewise linear space get by gluing together flat simplicial building blocks, which in dimension $d$ are $dD$ generalizations of triangles. By ‘flat’ we mean that they are isometric to a subspace of $dD$ Euclidean or Minkowski space. We will only be interested in gluings leading to genuine manifolds, which therefore look locally like an $R^d$. A nice feature of such simplicial manifolds is that their geometric properties are completely described by the discrete set $\{l_i^2\}$ of the squared lengths of their edges. Note that this amounts to a description of geometry without the use of coordinates. There is nothing to prevent us from re-introducing coordinate patches covering the piecewise linear manifold, for example, on each individual simplex, with suitable transition functions between patches. In such a coordinate system the metric tensor will then assume a definite form. However, for the purposes of formulating the path integral we will not be interested in doing this, but rather work with the edge lengths, which constitute a direct, regularized parametrization of the space $\text{Diff}(M)$ of geometries.

Fig. 3.8. Positive (a) and negative (b) space–like deficit angles $\delta$ (adapted from Loll (2001); Loll (1998)).
How precisely is the intrinsic geometry of a simplicial space, most importantly, its curvature, encoded in its edge lengths? A useful example to keep in mind is the case of dimension two, which can easily be visualized. A 2d piecewise linear space is a triangulation, and its scalar curvature $R(x)$ coincides with the Gaussian curvature (see [Ivancevic and Ivancevic 2007b]). One way of measuring this curvature is by parallel–transporting a vector around closed curves in the manifold. In our piecewise–flat manifold such a vector will always return to its original orientation unless it has surrounded lattice vertices $v$ at which the surrounding angles did not add up to $2\pi$, but $\sum_{i \supset v} \alpha_i = 2\pi - \delta$, for $\delta \neq 0$, see Figure 3.8. The so–called deficit angle $\delta$ is precisely the rotation angle picked up by the vector and is a direct measure for the scalar curvature at the vertex. The operational description to get the scalar curvature in higher dimensions is very similar, one basically has to sum in each point over the Gaussian curvatures of all 2D submanifolds. This explains why in Regge calculus the curvature part of the EH action is given by a sum over building blocks of dimension $(d-2)$ which are the objects dual to those local 2D submanifolds. More precisely, the continuum curvature and volume terms of the action become

$$\frac{1}{2} \int_R d^d x \sqrt{\det g^{(d)}} \rightarrow \sum_{i \in R} \text{Vol}(i^{th} (d-2)-simplex) \delta_i \quad (3.27)$$

$$\int_R d^d x \sqrt{\det g} \rightarrow \sum_{i \in R} \text{Vol}(i^{th} d-simplex) \quad (3.28)$$

in the simplicial discretization. It is then a simple exercise in trigonometry to express the volumes and angles appearing in these formulas as functions of the edge lengths $l_i$, both in the Euclidean and the Minkowskian case.

The approach of dynamical triangulations uses a certain class of such simplicial space–times as an explicit, regularized realization of the space Diff$(M)$. For a given volume $N_d$, this class consists of all gluings of manifold–type of a set of $N_d$ simplicial building blocks of top–dimension $d$ whose edge lengths are restricted to take either one or one out of two values. In the Euclidean case we set $l_i^2 = a^2$ for all $i$, and in the Lorentzian case we allow for both space– and time–like links with $l_i^2 \in \{-a^2, a^2\}$, where the geodesic distance $a$ serves as a short-distance cutoff, which will be taken to zero later. Coming from the classical theory this may seem a grave restriction at first, but this is indeed not the case. Firstly, keep in mind that for the purposes of the quantum theory we want to sample the space of geometries ‘ergodically’ at a coarse-grained scale of order $a$. This should
be contrasted with the classical theory where the objective is usually to approximate a given, fixed space–time to within a length scale \( a \). In the latter case one typically requires a much finer topology on the space of metrics or geometries. It is also straightforward to see that no local curvature degrees of freedom are suppressed by fixing the edge lengths; deficit angles in all directions are still present, although they take on only a discretized set of values. In this sense, in dynamical triangulations all geometry is in the gluing of the fundamental building blocks. This is dual to how quantum Regge calculus is set up, where one usually fixes a triangulation \( T \) and then ‘scans’ the space of geometries by letting the \( l_i \)'s run continuously over all values compatible with the triangular inequalities.

In a nutshell, Lorentzian dynamical triangulations give a definite meaning to the ‘integral over geometries’, namely, as a sum over inequivalent Lorentzian gluings \( T \) over any number \( N_d \) of \( d \)–simplices,

\[
\int_{\text{Geom}(\mathcal{M})} D[g_{\mu\nu}] e^{iS[g_{\mu\nu}]} \xrightarrow{\text{LDT}} \sum_{T \in \mathcal{T}} \frac{1}{C_T} e^{iS^{\text{Reg}}(T)},
\]

where the symmetry factor \( C_T = |\text{Aut}(T)| \) on the r.h.s. is the order of the automorphism group of the triangulation, consisting of all maps of \( T \) onto itself which preserve the connectivity of the simplicial lattice. We will specify below what precise class \( T \) of triangulations should appear in the summation.

It follows from the above that in this formulation all curvatures and volumes contributing to the Regge simplicial action come in discrete units. This can be illustrated by the case of a 2D triangulation with Euclidean signature, which according to the prescription of dynamical triangulations consists of equilateral triangles with squared edge lengths \(+a^2\). All interior angles of such a triangle are equal to \( \pi/3 \), which implies that the deficit angle at any vertex \( v \) can take the values \( 2\pi - k_v \pi/3 \), where \( k_v \) is the number of triangles meeting at \( v \). As a consequence, the Einstein–Regge action \( S^{\text{Reg}} \) assumes the simple form

\[
S^{\text{Reg}}(T) = \kappa_{d-2} N_{d-2} - \kappa_d N_d,
\]

where the coupling constants \( \kappa_i = \kappa_i(\lambda, G_N) \) are simple functions of the bare cosmological and Newton constants in \( d \) dimensions. Substituting this into the path sum in (3.29) leads to

\[
Z(\kappa_{d-2}, \kappa_d) = \sum_{N_d} e^{-i\kappa_d N_d} \sum_{N_{d-2}} e^{i\kappa_{d-2} N_{d-2}} \sum_{T|N_d, N_{d-2}} \frac{1}{C_T},
\]
The point of taking separate sums over the numbers of $d-$ and $(d - 2)-$simplices in (3.31) is to make explicit that ‘doing the sum’ is tantamount to the combinatorial problem of counting triangulations of a given volume and number of simplices of codimension 2 (corresponding to the last summation in (3.31)).

It turns out that at least in two space–time dimensions the counting of geometries can be done completely explicitly, turning both Lorentzian and Euclidean quantum gravity into exactly solvable statistical models.

3.1.6.4 Lorentzian Path Integral

Now, the simplicial building blocks of the models are taken to be pieces of Minkowski space, and their edges have squared lengths $+a^2$ or $-a^2$. For example, the two types of 4–simplices that are used in Lorentzian dynamical triangulations in dimension four are shown in Figure 3.9. The first of them has four time–like and six space–like links (and therefore contains 4 time–like and 1 space–like tetrahedron), whereas the second one has six time–like and four space–like links (and contains 5 time–like tetrahedra). Since both are subspaces of flat space with signature $(-++++)$, they possess well–defined light–cone structures everywhere [Loll (2001); Loll (1998)].

In general, gluings between pairs of $d-$simplices are only possible when the metric properties of their $(d - 1)-$faces match. Having local light cones implies causal relations between pairs of points in local neighborhoods. Creating closed time–like curves will be avoided by requiring that all space–times contributing to the path sum possess a global ‘time’ function $t$. In terms of the triangulation this means that the $d-$simplices are arranged such that their space–like links all lie in slices of constant integer $t$, and their time–like links interpolate between adjacent spatial slices $t$ and $t + 1$. Moreover, with respect to this time, we will not allow for any spatial topology changes.

---

11 The symmetry factor $C_T$ is almost always equal to 1 for large triangulations.

12 Note that if we were in the continuum and had introduced coordinates on space–time,
This latter condition is always satisfied in classical applications, where ‘trouser points’ like the one depicted in Figure 3.10 are ruled out by the requirement of having a non–degenerate Lorentzian metric defined everywhere on $M$ (it is geometrically obvious that the light cone and hence $g_{\mu\nu}$ must degenerate in at least one point along the ‘crotch’). Another way of thinking about such configurations (and their time–reversed counterparts) is that the causal past (future) of an observer changes discontinuously as her world–line passes near the singular point (see [Dowker (2002)] and references therein for related discussions about the issue of topology change in quantum gravity).

There is no a priori reason in the quantum theory to not relax some of these classical causality constraints. After all, as we stressed right at the outset, path integral histories are not in general classical solutions, nor can we attribute any other direct physical meaning to them individually. It might well be that one can construct models whose path integral configurations violate causality in this strict sense, but where this notion is somehow recovered in the resulting continuum theory. What the approach of Lorentzian dynamical triangulations has demonstrated is that \textit{imposing causality constraints will in general lead to a different continuum theory}. This is in contrast with the intuition one may have that ‘including a few isolated singular points will not make any difference’. On the contrary, tampering with causality in this way is not innocent at all, as was already anticipated by Teitelboim many years ago [Teitelboim (1983)].

We want to point out that one cannot conclude from the above that spatial topology changes or even \textit{fluctuations in the space–time topology} cannot be treated in the formulation of dynamical triangulations. However, if one insists on including geometries of variable topology in a Lorentzian discrete

such a statement would actually be diffeomorphism–invariant.
context, one has to come up with a prescription of how to weigh these singular points in the path integral, both before and after the Wick rotation [Dasgupta (2002)]. Maybe this can be done along the lines suggested in [Louko and Sorkin (1997)]; this is clearly an interesting issue for further research.

Having said this, we next have to address the question of the Wick rotation, in other words, of how to get rid of the factor of \( i \) in the exponent of (3.31). Without it, this expression is an infinite sum (since the volume can become arbitrarily large) of complex terms whose convergence properties will be very difficult to establish. In this situation, a Wick rotation is simply a technical tool which – in the best of all worlds – enables us to perform the state sum and determine its continuum limit. The end result will have to be Wick–rotated back to Lorentzian signature.

Fortunately, Lorentzian dynamical triangulations come with a natural notion of Wick rotation, and the strategy we just outlined can be carried out explicitly in two space–time dimensions, leading to a unitary theory. In higher dimensions we do not yet have sufficient analytical control of the continuum theories to make specific statements about the inverse Wick rotation. Since we use the Wick rotation at an intermediate step, one can ask whether other Wick rotations would lead to the same result. Currently this is a somewhat academic question, since it is in practice difficult to find such alternatives. In fact, it is quite miraculous we have found a single prescription for Wick–rotating in our regularized setting, and it does not seem to have a direct continuum analogue (for more comments on this issue, see [Dasgupta and Loll (2001); Dasgupta (2002)].

Our Wick rotation \( W \) in any dimension is an injective map from Lorentzian– to Euclidean–signature simplicial space–times. Using the notation \( T \) for a simplicial manifold together with length assignments \( l_s^2 \) and \( l_t^2 \) to its space– and time–like links, it is defined by

\[
T_{\text{lor}} = (T, \{ l_s^2 = a^2, l_t^2 = -a^2 \}) \quad \mapsto \quad W \quad T_{\text{eu}} = (T, \{ l_s^2 = a^2, l_t^2 = a^2 \}). \tag{3.32}
\]

Note that we have not touched the connectivity of the simplicial manifold \( T \), but only its metric properties, by mapping all time–like links of \( T \) into space–like ones, resulting in a Euclidean ‘space–time’ of equilateral building blocks. It can be shown [Ambjørn et al. (2001a)] that at the level of the corresponding weight factors in the path integral this Wick rotation has

\[13\]To get a genuine Wick rotation and not just a discrete map, one introduces a complex parameter \( \alpha \) in \( l_t^2 = -\alpha a^2 \). The proper prescription leading to (3.33) is then an analytic continuation of \( \alpha \) from 1 to \(-1\) through the lower–half complex–plane.
precisely the desired effect of rotating to the exponentiated Regge action of the ‘Euclideanized’ geometry,

\[ e^{iS(T_{\text{hor}})} \xrightarrow{W} e^{-S(T_{\text{eu}})}. \]  

(3.33)

The Euclideanized path sum after the Wick rotation has the form

\[ Z_{\text{eu}}(\kappa_d-2, \kappa_d) = \sum_{T} \frac{1}{C_T} e^{-\kappa_d N_d(T) + \kappa_d - 2 N_d - 2(T)} \]

\[ = \sum_{N_d} e^{-\kappa_d N_d} \sum_{T|N_d} \frac{1}{C_T} e^{\kappa_d - 2 N_d - 2(T)} \]  

(3.34)

\[ = \sum_{N_d} e^{-\kappa_d N_d} e^{\kappa_d^{\text{crit}}(\kappa_d - 2) N_d} \times \text{subleading}(N_d). \]

In the last equality we have used that the number of Lorentzian triangulations of discrete volume \( N_d \) to leading order scales exponentially with \( N_d \) for large volumes. This can be shown explicitly in space–time dimension 2 and 3. For \( d = 4 \), there is strong (numerical) evidence for such an exponential bound for Euclidean triangulations, from which the desired result for the Lorentzian case follows (since \( W \) maps to a strict subset of all Euclidean simplicial manifolds).

From the functional form of the last line of (3.34) one can immediately read off some qualitative features of the phase diagram, an example of which appeared already earlier in Figure 3.7. Namely, the sum over geometries \( Z_{\text{eu}} \) converges for values \( \kappa_d > \kappa_d^{\text{crit}} \) of the bare cosmological constant, and diverges (i.e., is not defined) below this critical line. Generically, for all models of dynamical triangulations the infinite–volume limit is attained by approaching the critical line \( \kappa_d^{\text{crit}}(\kappa_d - 2) \) from above, i.e. from inside the region of convergence of \( Z_{\text{eu}} \). In the process of taking \( N_d \to \infty \) and the cutoff \( a \to 0 \), one gets a renormalized cosmological constant \( \Lambda \) through

\[ (\kappa_d - \kappa_d^{\text{crit}}) = a^{d-1} \Lambda + O(a^{d+1}). \]  

(3.35)

If the scaling is canonical (which means that the dimensionality of the renormalized coupling constant is the one expected from the classical theory), the exponent is given by \( \mu = d \). Note that this construction requires a positive bare cosmological constant in order to make the state sum converge. Moreover, by virtue of relation (3.35) also the renormalized cosmological constant must be positive. Other than that, its numerical value is not determined by this argument, but by comparing observables of the theory which depend
on $A$ with actual physical measurements. Another interesting observation is that the inclusion of a sum over topologies in the discretized sum (3.34) would lead to a super–exponential growth of at least $\propto N_d!$ of the number of triangulations with the volume $N_d$. Such a divergence of the path integral cannot be compensated by an additive renormalization of the cosmological constant of the kind outlined above.

There are ways in which one can sum divergent series of this type, for example, by performing a Borel sum. The problem with these stems from the fact that two different functions can share the same asymptotic expansion. Therefore, the series in itself is not sufficient to define the underlying theory uniquely. The non–uniqueness arises because of non–perturbative contributions to the path integral which are not represented in the perturbative expansion. In order to fix these uniquely, an independent, non–perturbative definition of the theory is necessary. Unfortunately, for dynamically triangulated models of quantum gravity, no such definitions have been found so far. In the context of 2D (Euclidean) quantum gravity this difficulty is known as the ‘absence of a physically motivated double-scaling limit’ [Ambjørn and Kristjansen (1993)].

Lastly, getting an interesting continuum limit may or may not require an additional fine–tuning of the inverse gravitational coupling $\kappa_{d-2}$, depending on the dimension $d$. In four dimensions, one would expect to find a second-order transition along the critical line, corresponding to local gravitonic excitations. The situation in $d=3$ is less clear, but results get so far indicate that no fine–tuning of Newton’s constant is necessary [Ambjørn et al. (2001b); Ambjørn et al. (2001d)].

Before delving into the details, let me summarize briefly the results that have been get so far in the approach of Lorentzian dynamical triangulations. At the regularized level, that is, in the presence of a finite cutoff $a$ for the edge lengths and an infrared cutoff for large space–time volume, they are well–defined statistical models of Lorentzian random geometries in $d=2,3,4$. In particular, they obey a suitable notion of reflection-positivity and possess self–adjoint Hamiltonians.

The crucial questions are then to what extent the underlying combinatorial problems of counting all $dD$ geometries with certain causal properties can be solved, whether continuum theories with non–trivial dynamics ex-
quantum leap

3.2 Dynamics of Quantum Fields

3.2.1 Path Integrals and Green’s Functions

In quantum field theory we are interested in objects such as \( \langle 0 | T^{\hat{\phi}(x_1)\hat{\phi}(x_2)\cdots\hat{\phi}(x_n)} | 0 \rangle \),
the vacuum expectation value of a time-ordered product of Heisenberg field operators. This object is known as a Green’s function, or as a correlation

\[ \langle 0 | T^{\hat{\phi}(x_1)\hat{\phi}(x_2)\cdots\hat{\phi}(x_n)} | 0 \rangle , \]

Recall that a Green’s function, \( G(x,s) \), of a linear operator \( L \) acting on distributions over a manifold \( M \), at a point \( x_0 \), is any solution of linear operator equation

\[ LG(x,s) = \delta(x-s) , \]

where \( \delta(x) \) is the Dirac delta function at a point \( x \in M \). This technique can be used to solve differential equations of the form

\[ Lu(x) = f(x) . \]

If the kernel of the operator \( L \) is nontrivial, then the Green’s function is not unique. However, in practice, some combination of symmetry, boundary conditions and/or other externally imposed criteria would give us a unique Green’s function. Also, Green’s functions in general are distributions (or generalized functions like the Dirac delta function \( \delta(x) \)), not necessarily proper functions.

Green’s functions are a useful tool in condensed matter theory, where they allow the resolution of the diffusion equation, as well as in quantum mechanics, where the Green’s function of the Hamiltonian is a key concept, with important links to the concept of density of states. The Green’s functions used in those two domains are highly similar,
function. The order of the operators is such that the earliest field is written last (right-most), the second earliest second last, etc. For example,

\[ T\hat{\phi}(x_1)\hat{\phi}(x_2) = \begin{cases} \hat{\phi}(x_1)\hat{\phi}(x_2), & x_1^0 > x_2^0, \\ \hat{\phi}(x_2)\hat{\phi}(x_1), & x_2^0 > x_1^0. \end{cases} \]

Green’s functions are related to amplitudes for physical processes such as scattering and decay processes.

Let us look at the analogous object in quantum mechanics,

\[ G^{(n)}(t_1, t_2, \ldots, t_n) = \langle 0 | T\hat{q}(t_1)\hat{q}(t_2) \cdots \hat{q}(t_n) | 0 \rangle. \]

We will develop a path–integral expression for this.

First, we must recast the path–integral in terms of Heisenberg representation objects. The operator \( \hat{q}(t) \) is the usual Heisenberg operator, defined due to the analogy in the mathematical structure of the real–valued diffusion equation and complex–valued Schrödinger equation. Briefly, if such a function \( G(x, s) \) can be found for the operator \( L \), then if we multiply the equation \( LG(x, s) = \delta(x - s) \) for the Green's function by \( f(s) \), and then perform an integration in the \( s \) variable, we get

\[ \int LG(x, s)f(s)ds = \int \delta(x - s)f(s)ds = f(x). \]

As the right hand side is actually \( Lu(x) = f(x) \), we get

\[ Lu(x) = \int LG(x, s)f(s)ds. \]

Further, because the operator \( L \) is linear and acts on the variable \( x \) alone (not on the variable of integration \( s \)), we can take the operator \( L \) outside of the integration on the right hand side, getting

\[ Lu(x) = L \left( \int G(x, s)f(s)ds \right), \]

which implies

\[ u(x) = \int G(x, s)f(s)ds. \]

Thus, we can obtain the function \( u(x) \) through knowledge of the Green's function in the first equation, and the source term on the right hand side in second equation. This process has resulted from the linearity of the operator \( L \).

A convolution with a Green’s function gives solutions to inhomogeneous differential–integral equations, most commonly a Sturm–Liouville problem. If \( G \) is the Green’s function of an operator \( L \), then the solution for \( u \) of the equation \( Lu = f \) is given by

\[ u(x) = \int f(s)G(x, s)ds. \]

This can be thought of as an expansion of the function \( f \) according to a Dirac delta function basis (projecting \( f \) over \( \delta(x - s) \)) and a superposition of the solution on each projection. Such an integral is known as a Fredholm integral equation, the study of which constitutes Fredholm theory.
in terms of the Schrödinger operator $\hat{q}$ by

$$\hat{q}(t) = e^{iHt} \hat{q} e^{-iHt}.$$ 

The eigenstates of the Heisenberg operator are $|q, t\rangle$: $\hat{q}(t) |q, t\rangle = q |q, t\rangle$.

The relation with the time–independent eigenstates is $|q, t\rangle = e^{iHt} |q\rangle$ [MacKenzie (2000)]. Then we can write the path–integral,

$$K = \langle q' | e^{-iHT} | q \rangle = \int D[q] e^{iS}.$$ 

We can now calculate the two–point function $G(t_1, t_2)$, via the path–integral. We will proceed in two steps. First, we will calculate the following expression:

$$\langle q', T | T\hat{q}(t_1)\hat{q}(t_2) | q, 0 \rangle.$$ 

We will then devise a method for extracting the vacuum contribution to the initial and final states.

Suppose first that $t_1 > t_2$. Then

$$\langle q', T | T\hat{q}(t_1)\hat{q}(t_2) | q, 0 \rangle = \langle q', T | \hat{q}(t_1)\hat{q}(t_2) | q, 0 \rangle$$

$$= \int dq_1 dq_2 \langle q', T | q_1, t_1 \rangle \langle q_1, t_1 | \hat{q}(t_1)\hat{q}(t_2) | q_2, t_2 \rangle \langle q_2, t_2 | q, 0 \rangle$$

$$= \int dq_1 dq_2 q_1 q_2 \langle q', T | q_1, t_1 \rangle \langle q_1, t_1 | q_2, t_2 \rangle \langle q_2, t_2 | q, 0 \rangle.$$ 

Each of these matrix elements is a path–integral:

$$\langle q', T | T\hat{q}(t_1)\hat{q}(t_2) | q, 0 \rangle = \int dq_1 dq_2 \int_{q_1, t_1}^{q', T} D[q] e^{iS} \int_{q_2, t_2}^{q_1, t_1} D[q] e^{iS} \int_{q, 0}^{q_2, t_2} D[q] e^{iS}.$$ 

This expression consists of a first path–integral from the initial position $q$ to an arbitrary position $q_2$, a second one from there to a second arbitrary position $q_1$, and a third one from there to the final position $q'$. So we are integrating over all paths from $q$ to $q'$, subject to the restriction that the paths pass through the intermediate points $q_1$ and $q_2$. We then integrate over the two arbitrary positions, so that in fact we are integrating over all paths: we can combine these three path integrals plus the integrations over $q_1$ and $q_2$ into one path–integral. The factors $q_1$ and $q_2$ in the above
integral can be incorporated into this path–integral by simply including a factor \( q(t_1)q(t_2) \) in the path–integral. So

\[
\langle q', T | \hat{q}(t_1)\hat{q}(t_2) | q, 0 \rangle = \int_{q,0}^{q',T} D[q] q(t_1)q(t_2)e^{iS}, \quad (t_1 > t_2).
\]

An identical calculation shows that exactly this same final expression is also valid for \( t_2 < t_1 \): magically, the path–integral does the time ordering automatically. Thus for all times

\[
\langle q', T | T\hat{q}(t_1)\hat{q}(t_2) \ldots \hat{q}(t_n) | q, 0 \rangle = \int_{q,0}^{q',T} D[q] q(t_1)q(t_2)\ldots q(t_n)e^{iS}.
\]

As for how to obtain vacuum-to-vacuum matrix elements, our work on statistical mechanics provides us with a clue. We can expand the states \( \langle q', T \rangle \) and \( \langle q, 0 \rangle \) in terms of eigenstates of the Hamiltonian. If we evolve towards a negative imaginary time, the contribution of all other states will decay away relative to that of the ground state. We have (resetting the initial time to \( -T \) for convenience)

\[
\langle 0, T | T\hat{q}(t_1)\hat{q}(t_2) \ldots \hat{q}(t_n) | 0, -T \rangle \propto \langle 0, T| q, -T \rangle,
\]

where on the right the ‘0’ denotes the ground state. The proportionality involves the ground state wave function and an exponential factor

\[
\exp[2iE_0T] = \exp[-2E_0|T|].
\]

We could perform all calculations in a Euclidean theory and analytically continue to real time when computing physical quantities (many books do this), but to be closer to physics we can also consider \( T \) not to be pure imaginary and negative, but to have a small negative imaginary phase: \( T = |T|e^{-i\epsilon} \) (\( \epsilon > 0 \)). With this,

\[
\langle 0, T | q, -T \rangle \propto \langle q', T | q, -T \rangle = \int D[q] e^{iS}.
\]

To compute the Green’s functions, we must simply add \( T\hat{q}(t_1)\hat{q}(t_2) \ldots \hat{q}(t_n) \) to the matrix element, and the corresponding factor \( q(t_1)q(t_2)\ldots q(t_n) \) inside the path–integral:

\[
\langle 0, T | T\hat{q}(t_1)\hat{q}(t_2) \ldots \hat{q}(t_n) | 0, -T \rangle \propto \int D[q] q(t_1)q(t_2)\ldots q(t_n)e^{iS}.
\]

The proportionality sign is a bit awkward; fortunately, we can rid ourselves of it. To do this, we note that the left hand expression is not exactly what
we want: the vacua $|0, \pm T\rangle$ differ by a phase. We wish to eliminate this phase; to this end, the Green’s functions are defined

$$G^{(n)}(t_1, t_2, \cdots, t_n) = \frac{\langle 0 | T \hat{q}(t_1) \hat{q}(t_2) \cdots \hat{q}(t_n) | 0 \rangle}{\langle 0, T | 0, -T \rangle} = \frac{\int \mathcal{D}[q] q(t_1) q(t_2) \cdots q(t_n) e^{iS}}{\int \mathcal{D}[q] e^{iS}},$$

with no proportionality sign. The wave functions and exponential factors in the numerator and denominator cancel.

To compute the numerator, we can once again use the trick we used in perturbation theory in quantum mechanics, namely, adding a source to the action. We define

$$Z[J] = \frac{\int \mathcal{D}[q] e^{i(S + \int dt J(t) \hat{q}(t))}}{\int \mathcal{D}[q] e^{iS}} = \frac{\langle 0 | 0 \rangle_J}{\langle 0 \rangle_{J = 0}}.$$  

If we operate on $Z[J]$ with $i^{-1} \delta/\delta J(t_1)$, this gives

$$\left. \left( \frac{1}{i} \frac{\delta}{\delta J(t_1)} Z[J] \right) \right|_{J = 0} = \left. \left( \frac{\int \mathcal{D}[q] q(t_1) e^{i(S + \int dt J(t) \hat{q}(t))}}{\int \mathcal{D}[q] e^{iS}} \right) \right|_{J = 0} = \frac{\langle 0, T | \hat{q}(t_1) | 0, -T \rangle}{\langle 0, T | 0, -T \rangle} = \langle 0 | \hat{q}(t_1) | 0 \rangle.$$

(The expectation values are evaluated in the absence of $J$.)

Repeating this procedure, we obtain a path–integral with several $q$’s in the numerator. This ordinary product of $q$’s in the path–integral corresponds, as discussed earlier in this section, to a time-ordered product in the matrix element. So we make the following conclusion:

$$\left. \left( \frac{1}{i} \frac{\delta}{\delta J(t_1)} \cdots \frac{1}{i} \frac{\delta}{\delta J(t_n)} Z[J] \right) \right|_{J = 0} = \frac{\int \mathcal{D}[q] q(t_1) \cdots q(t_n) e^{iS}}{\int \mathcal{D}[q] e^{iS}} = \langle 0 | T \hat{q}(t_1) \cdots \hat{q}(t_n) | 0 \rangle.$$

For obvious reasons, the functional $Z[J]$ is called the generating functional for Green’s functions; it is a very handy tool in quantum field theory and in statistical mechanics.

To be able to calculate $Z[J]$, let us examine the numerator,

$$N \equiv \int \mathcal{D}[q] e^{i(S + \int dt J(t) \hat{q}(t))}.$$  

Suppose initially that $S$ is the harmonic oscillator action (denoted $S_0$):

$$S_0 = \int dt \left( \frac{1}{2} m \dot{q}^2 - \frac{1}{2} m \omega^2 q^2 \right),$$

Then the corresponding numerator, $N_0$, is the non–Euclidean, or real-time, version of the propagator $K_0^E[J]$ that we used before. We can calculate $N_0[J]$ in the same way as $K_0^E[J]$. Since the calculation repeats much of that of $K_0^E[J]$, we will be succinct.

By definition,

$$N_0 = \int Dq(t) \exp \left[ i \int dt \left( \frac{1}{2} m \dot{q}^2 - \frac{1}{2} m \omega^2 q^2 + Jq \right) \right].$$

We do the path integral over a new variable $y$, defined by $q(t) = q_c(t) + y(t)$, where $q_c$ is the classical solution. Then the path–integral over $y$ is a constant (independent of $J$) and we can avoid calculating it. (It will cancel against the denominator in $Z[J]$.) Calling it $C$, we have

$$N_0 = C e^{i S_0[J]}.$$

We can now write

$$N_0 = C \exp \left[ \frac{1}{2} \int dt \, dt' \,\dot{J}(t) G(t,t') J(t') \right].$$

Dividing by the denominator merely cancels the factor $C$, giving our final result [MacKenzie (2000)]:

$$Z[J] = \exp \left[ \frac{1}{2} \int dt \, dt' \,\dot{J}(t) G(t,t') J(t') \right].$$
We can solve (3.36) for the Green’s function by going into momentum space; the result is

\[ G(t, t') = \int \frac{dk}{2\pi} \frac{i}{k^2 - \omega^2} e^{-ik(t-t')} . \]

However, there are poles on the axis of integration. The Green’s function is ambiguous until we give it a ‘pole prescription’, i.e., a boundary condition. But remember that our time \( T \) has a small, negative imaginary part. We require that \( G \) go to zero as \( T \to \infty \). The correct pole prescription then turns out to be

\[ G(t-t') = \int \frac{dk}{2\pi} \frac{i}{k^2 - \omega^2 + i\epsilon} e^{-ik(t-t')} . \] (3.37)

We could at this point do a couple of practice calculations to get used to this formalism. Examples would be to compute perturbatively the generating functional for an action which has terms beyond quadratic (for example, a \( q^4 \) term), or to compute some Green’s function in either the quadratic or quartic theory. But since these objects are not really useful in quantum mechanics, without further delay we will go directly to the case of interest: quantum field theory.

### 3.2.2 Topological Quantum Field Theory

Before we come to (super)strings, we give a brief on topological quantum field theory (TQFT), as developed by Ed Witten, from his original path integral point of view (see \[ Witten (1988b) \] Labastida and Lozano (1998)). TQFT originated in 1982, when Witten rewrote classical Morse theory (see Ivanovic and Ivanovic (2007b)) in Dick Feynman’s language of quantum field theory \[ Witten (1982) \]. Witten’s arguments made use of Feynman’s path integrals and consequently, at first, they were regarded as mathematically non–rigorous. However, a few years later, A. Floer reformulated a rigorous Morse–Witten theory \[ Floer (1987) \] (that won a Fields medal for Witten). This trend in which some mathematical structure is first constructed by quantum field theory methods and then reformulated in a rigorous mathematical ground constitutes one of the tendencies in modern physics.

In TQFT our basic topological space is an \( n \)D Riemannian manifold \( M \) with a metric \( g_{\mu\nu} \). Let us consider on it a set of fields \( \{ \phi_i \} \), and let \( S[\phi_i] \) be a real functional of these fields which is regarded as the action of the theory. We consider ‘operators’, \( O_\alpha(\phi_i) \), which are in general arbitrary
functionals of the fields. In TQFT these functionals are real functionals labelled by some set of indices $\alpha$ carrying topological or group–theoretical data. The vacuum expectation value (VEV) of a product of these operators is defined as

$$\langle O_{\alpha_1} O_{\alpha_2} \cdots O_{\alpha_p} \rangle = \int [D\phi_i] O_{\alpha_1}(\phi_i) O_{\alpha_2}(\phi_i) \cdots O_{\alpha_p}(\phi_i) \exp(-S[\phi_i]).$$

A quantum field theory is considered topological if the following relation is satisfied:

$$\frac{\delta}{\delta g_{\mu\nu}} (O_{\alpha_1} O_{\alpha_2} \cdots O_{\alpha_p}) = 0,$$

i.e., if the VEV of some set of selected operators is independent of the metric $g_{\mu\nu}$ on $M$. If such is the case those operators are called ‘observables’.

There are two ways to guarantee, at least formally, that condition (3.38) is satisfied. The first one corresponds to the situation in which both, the action $S[\phi_i]$, as well as the operators $O_{\alpha_i}$ are metric independent. These TQFTs are called of Schwarz type. The most important representative is Chern–Simons gauge theory. The second one corresponds to the case in which there exist a symmetry, whose infinitesimal form is denoted by $\delta$, satisfying the following properties:

$$\delta O_{\alpha_i} = 0, \quad T_{\mu\nu} = \delta G_{\mu\nu},$$

where $T_{\mu\nu}$ is the SEM–tensor of the theory, i.e.,

$$T_{\mu\nu}(\phi_i) = \frac{\delta}{\delta g_{\mu\nu}} S[\phi_i].$$

The fact that $\delta$ in (3.39) is a symmetry of the theory implies that the transformations $\delta \phi_i$ of the fields are such that both $\delta A[\phi_i] = 0$ and $\delta O_{\alpha_i}(\phi_i) = 0$. Conditions (3.39) lead, at least formally, to the following relation for VEVs:

$$\frac{\delta}{\delta g_{\mu\nu}} (O_{\alpha_1} O_{\alpha_2} \cdots O_{\alpha_p}) = - \int [D\phi_i] O_{\alpha_1}(\phi_i) O_{\alpha_2}(\phi_i) \cdots O_{\alpha_p}(\phi_i) T_{\mu\nu} e^{-S[\phi_i]}$$

$$= - \int [D\phi_i] \delta (O_{\alpha_1}(\phi_i) O_{\alpha_2}(\phi_i) \cdots O_{\alpha_p}(\phi_i)) G_{\mu\nu} \exp(-S[\phi_i])) = 0,$$

which implies that the quantum field theory can be regarded as topological. This second type of TQFTs are called of Witten type. One of its main representatives is the theory related to Donaldson invariants, which is a twisted version of $N = 2$ supersymmetric Yang–Mills gauge theory. It
Quantum Leap

is important to remark that the symmetry $\delta$ must be a scalar symmetry, i.e., that its symmetry parameter must be a scalar. The reason is that, being a global symmetry, this parameter must be covariantly constant and for arbitrary manifolds this property, if it is satisfied at all, implies strong restrictions unless the parameter is a scalar.

Most of the TQFTs of cohomological type satisfy the relation:

$$S[\phi_i] = \delta \Lambda(\phi_i),$$

(3.42)

for some functional $\Lambda(\phi_i)$. This has far-reaching consequences, for it means that the topological observables of the theory, in particular the partition function, (path integral) itself are independent of the value of the coupling constant. Indeed, let us consider for example the VEV:

$$\langle O_{\alpha_1}O_{\alpha_2}\cdots O_{\alpha_p} \rangle = \int [D\phi] O_{\alpha_1}(\phi_i)O_{\alpha_2}(\phi_i)\cdots O_{\alpha_p}(\phi_i) e^{-\frac{1}{g^2}S[\phi_i]}. \quad (3.43)$$

Under a change in the coupling constant, $1/g^2 \rightarrow 1/g^2 - \Delta$, one has (assuming that the observables do not depend on the coupling), up to first-order in $\Delta$:

$$\langle O_{\alpha_1}O_{\alpha_2}\cdots O_{\alpha_p} \rangle \rightarrow \langle O_{\alpha_1}O_{\alpha_2}\cdots O_{\alpha_p} \rangle + \Delta \int [D\phi] \delta \left[ O_{\alpha_1}(\phi_i)O_{\alpha_2}(\phi_i)\cdots O_{\alpha_p}(\phi_i)\Lambda(\phi) \right] e^{-\frac{1}{g^2}S[\phi_i]} \exp \left( -\frac{1}{g^2}S[\phi_i] \right). \quad (3.44)$$

Hence, observables can be computed either in the weak coupling limit, $g \rightarrow 0$, or in the strong coupling limit, $g \rightarrow \infty$.

So far we have presented a rather general definition of TQFT and made a series of elementary remarks. Now we will analyze some aspects of its structure. We begin pointing out that given a theory in which (3.39) holds one can build correlators which correspond to topological invariants (in the sense that they are invariant under deformations of the metric $g_{\mu\nu}$) just by considering the operators of the theory which are invariant under the symmetry. We will call these operators observables. In virtue of (3.41), if one of these operators can be written as a symmetry transformation of another operator, its presence in a correlation function will make it vanish. Thus we may identify operators satisfying (3.39) which differ by an operator which corresponds to a symmetry transformation of another operator. Let us denote the set of the resulting classes by $\{\Phi\}$. By restricting the analysis
to the appropriate set of operators, one has that in fact,

$$\delta^2 = 0.$$ \hspace{1cm} (3.45)

Property (3.45) has consequences on the features of TQFT. First, the symmetry must be odd which implies the presence in the theory of commuting and anticommuting fields. For example, the tensor $G_{\mu\nu}$ in (3.39) must be anticommuting. This is the first appearance of an odd non-spinorial field in TQFT. Those kinds of objects are standard features of cohomological TQFTs. Second, if we denote by $Q$ the operator which implements this symmetry, the observables of the theory can be described as the cohomology classes of $Q$:

$$\{\Phi\} = \frac{\text{Ker} \, Q}{\text{Im} \, Q}, \quad Q^2 = 0.$$ \hspace{1cm} (3.46)

Equation (3.39) means that in addition to the Poincaré group the theory possesses a symmetry generated by an odd version of the Poincaré group. The corresponding odd generators are constructed out of the tensor $G_{\mu\nu}$ in much the same way as the ordinary Poincaré generators are built out of $T_{\mu\nu}$. For example, if $P_{\mu}$ represents the ordinary momentum operator, there exists a corresponding odd one $G_{\mu}$ such that

$$P_{\mu} = \{Q, G_{\mu}\}.$$ \hspace{1cm} (3.47)

Now, let us discuss the structure of the Hilbert space of the theory in virtue of the symmetries that we have just described. The states of this space must correspond to representations of the algebra generated by the operators in the Poincaré groups and by $Q$. Furthermore, as follows from our analysis of operators leading to (3.46), if one is interested only in states $|\Psi\rangle$ leading to topological invariants one must consider states which satisfy

$$Q|\Psi\rangle = 0,$$ \hspace{1cm} (3.48)

and two states which differ by a $Q$–exact state must be identified. The odd Poincaré group can be used to generate descendant states out of a state satisfying (3.48). The operators $G_{\mu}$ act non-trivially on the states and in fact, out of a state satisfying (3.48) we can build additional states using this generator. The simplest case consists of

$$\int_{\gamma_1} G_{\mu} |\Psi\rangle,$$
where $\gamma_1$ is a 1–cycle. One can verify using (3.39) that this new state satisfies (3.48):

$$Q \int_{\gamma_1} G_\mu |\Psi\rangle = \int_{\gamma_1} \{Q, G_\mu\} |\Psi\rangle = \int_{\gamma_1} P_\mu |\Psi\rangle = 0.$$  

Similarly, one may construct other invariants tensoring $n$ operators $G_\mu$ and integrating over $n$–cycles $\gamma_n$:

$$\int_{\gamma_n} G_{\mu_1} G_{\mu_2} ... G_{\mu_n} |\Psi\rangle.$$  

(3.49)

Notice that since the operator $G_\mu$ is odd and its algebra is Poincaré–like the integrand in this expression is an exterior differential $n$–form. These states also satisfy condition (3.48). Therefore, starting from a state $|\Psi\rangle \in \ker Q$ we have built a set of partners or descendants giving rise to a topological multiplet. The members of a multiplet have well defined ghost number. If one assigns ghost number $-1$ to the operator $G_\mu$ the state in (3.49) has ghost number $-n$ plus the ghost number of $|\Psi\rangle$. Now, $n$ is bounded by the dimension of the manifold $X$. Among the states constructed in this way there may be many which are related via another state which is $Q$–exact, i.e., which can be written as $Q$ acting on some other state. Let us try to single out representatives at each level of ghost number in a given topological multiplet.

Consider an $(n-1)$–cycle which is the boundary of an $n$D surface, $\gamma_{n-1} = \partial S_n$. If one builds a state taking such a cycle one finds ($P_\mu = -i\partial_\mu$),

$$\int_{\gamma_{n-1}} G_{\mu_1} G_{\mu_2} ... G_{\mu_{n-1}} |\Psi\rangle = i \int_{S_n} P_{[\mu_1} G_{\mu_2} G_{\mu_3} ... G_{\mu_n]} |\Psi\rangle$$  

(3.50)

$$= iQ \int_{S_n} G_{\mu_1} G_{\mu_2} ... G_{\mu_n} |\Psi\rangle,$$

i.e., it is $Q$–exact. The square–bracketed subscripts in (3.50) denote that all indices between them must by antisymmetrized. In (3.50) use has been made of (3.47). This result tells us that the representatives we are looking for are built out of the homology cycles of the manifold $X$. Given a manifold $X$, the homology cycles are equivalence classes among cycles, the equivalence relation being that two $n$–cycles are equivalent if they differ by a cycle which is the boundary of an $n+1$ surface. Thus, knowledge on the homology of the manifold on which the TQFT is defined allows us to classify the representatives among the operators (3.49). Let us assume that $X$ has dimension $d$ and that its homology cycles are $\gamma_{i_n}$, ($i_n = 1, ..., d_n$,
$n = 0, \ldots, d$, where $d_n$ is the dimension of the $n$–homology group, and $d$ the dimension of $X$. Then, the non–trivial partners or descendants of a given $|\Psi\rangle$ highest–ghost–number state are labelled in the following way:

$$\int_{\gamma_n} G_{\mu_1} G_{\mu_2} \cdots G_{\mu_n} |\Psi\rangle, \quad (i_n = 1, \ldots, d_n, \ n = 0, \ldots, d).$$

A similar construction to the one just described can be made for fields. Starting with a field $\phi(x)$ which satisfies,

$$[Q, \phi(x)] = 0,$$

one can construct other fields using the operators $G_{\mu}$. These fields, which we call partners are antisymmetric tensors defined as,

$$\phi^{(n)}_{\mu_1 \mu_2 \cdots \mu_n}(x) = \frac{1}{n!} [G_{\mu_1}, [G_{\mu_2} \cdots [G_{\mu_n}, \phi(x)] \cdots]], \quad (n = 1, \ldots, d).$$

Using (3.47) and (3.51) one finds that these fields satisfy the so–called topological descent equations:

$$d\phi^{(n)} = i[Q, \phi^{(n+1)}],$$

where the subindices of the forms have been suppressed for simplicity, and the highest–ghost–number field $\phi(x)$ has been denoted as $\phi^{(0)}(x)$. These equations enclose all the relevant properties of the observables which are constructed out of them. They constitute a very useful tool to build the observables of the theory.

### 3.2.3 TQFT and Seiberg–Witten Theory

in the study of the Seiberg-Witten (SW) monopole equations, which have their origin in supersymmetric gauge theory. The SW theory, while closely related to Donaldson theory, is much easier to handle. Using SW theory, proofs of many theorems in Donaldson theory have been simplified, and several important new results have also been obtained [Taubes (1990), Taubes (1994)].

In [Zhang et al. (1995)] a topological quantum field theory was introduced which reproduces the SW invariants of 4-manifolds. A geometrical interpretation of the 3D quantum field theory was also given.

3.2.3.1 SW Invariants and Monopole Equations

Recall that the SW monopole equations are classical field theoretical equations involving a $U(1)$ gauge field and a complex Weyl spinor on a 4D manifold. Let $X$ denote the 4-manifold, which is assumed to be oriented and closed. If $X$ is spin, there exist positive and negative spin bundles $S^\pm$ of rank two. Introduce a complex line bundle $L \to X$. Let $A$ be a connection on $L$ and $M$ be a section of the product bundle $S^+ \otimes L$. Recall that the SW monopole equations read

$$ F^+_{kl} = -\frac{i}{2} \bar{M} \Gamma_{kl} M, \quad D_A M = 0, \quad (3.52) $$

where $D_A$ is the twisted Dirac operator, $\Gamma_{ij} = \frac{1}{2}[\gamma_i, \gamma_j]$, and $F^+$ represents the self-dual part of the curvature of $L$ with connection $A$.

If $X$ is not a spin manifold, then spin bundles do not exist. However, it is always possible to introduce the so called $\text{Spin}_c$ bundles $S^\pm \otimes L$, with $L^2$ being a line bundle. Then in this more general setting, the SW monopoles equations look formally the same as (3.52), but the $M$ should be interpreted as a section of the $\text{Spin}_c$ bundle $S^+ \otimes L$.

Denote by $M$ the moduli space of solutions of the SW monopole equations up to gauge transformations. Generically, this space is a manifold. Its virtual dimension is equal to the number of solutions of the following equations

$$ (d\psi)^+_{kl} + \frac{i}{2} (\bar{M} \Gamma_{kl} N + \bar{N} \Gamma_{kl} M) = 0, \quad D_A N + \psi M = 0, $$

$$ \nabla_k \psi^k + \frac{i}{2} (\overline{N} M - \overline{M} N) = 0, \quad (3.53) $$

where $A$ and $M$ are a given solution of (3.52), $\psi \in \Omega^1(X)$ is a one form, $(d\psi)^+ \in \Omega^2^+(X)$ is the self dual part of the two form $d\psi$, and $N \in S^+ \otimes L$. 

The first two of the equations in (3.53) are the linearization of the monopole equations (3.52), while the last one is a gauge fixing condition. Though with a rather unusual form, it arises naturally from the dual operator governing gauge transformations

\[ C : \Omega^0(X) \rightarrow \Omega^1(X) \oplus (S^+ \otimes L), \quad \phi \mapsto (-d\phi, i\phi M). \]

Let \( T : \Omega^1(X) \oplus (S^+ \otimes L) \rightarrow \Omega^0(X) \oplus \Omega^{2+}(X) \oplus (S^- \otimes L), \) be the operator governing equation (3.53), namely, the operator which allows us to rewrite (3.53) as

\[ T(\psi, N) = 0. \]

Then \( T \) is an elliptic operator, the index \( \text{Ind}(T) \) of which yields the virtual dimension of \( M \). A straightforward application of the Atiyah–Singer index Theorem gives

\[ \text{Ind}(T) = -\frac{2\chi(X) + 3\sigma(X)}{4} + c_1(L)^2, \]

where \( \chi(X) \) is the Euler character of \( X \), \( \sigma(X) \) its signature index and \( c_1(L)^2 \) is the square of the first Chern class of \( L \) evaluated on \( X \) in the standard way.

When \( \text{Ind}(T) \) equals zero, the moduli space generically consists of a finite number of points, \( M = \{ p_t : t = 1, 2, ..., I \} \). Let \( \epsilon_t \) denote the sign of the determinant of the operator \( T \) at \( p_t \), which can be defined with mathematical rigor. Then the SW invariant of the 4–manifold \( X \) is defined by

\[ \sum_{i=1}^{I} \epsilon_i. \]

The fact that this is indeed an invariant(i.e., independent of the metric) of \( X \) is not very difficult to prove, and we refer to [Witten (1994)] for details. As a matter of fact, the number of solutions of a system of equations weighted by the sign of the operator governing the equations(i.e., the analog of \( T \)) is a topological invariant in general [Witten (1994)]. This point of view has been extensively explored by Vafa and Witten [Vafa and Witten (1994)] within the framework of topological quantum field theory in connection with the so called S duality. Here we wish to explore the SW invariants following a similar line as that taken in [Witten (1988b); Vafa and Witten (1994)].

### 3.2.3.2 Topological Lagrangian

Introduce a Lie super–algebra with an odd generator \( Q \) and two even generators \( U \) and \( \delta \) obeying the following (anti)commutation relations [Zhang et al. (1995)]

\[ [U, Q] = Q, \quad [Q, Q] = 2\delta, \quad [Q, \delta] = 0. \]  

(3.54)
We will call $U$ the ghost number operator, and $Q$ the BRST–operator.

Let $A$ be a connection of $L$ and $M \in S^+ \otimes L$. We define the action of the super–algebra on these fields by requiring that $\delta$ coincide with a gauge transformation with a gauge parameter $\phi \in \Omega^0(X)$. The field multiplets associated with $A$ and $M$ furnishing representations of the super–algebra are $(A, \psi, \phi)$, and $(M,N)$, where $\psi \in \Omega^1(X)$, $\phi \in \Omega^0(X)$, and $N$ is a section of $S^+ \otimes L$. They transform under the action of the super–algebra according to

\[
\begin{align*}
&[Q,A_i] = \psi_i, \quad [Q,M] = N, \\
&[Q,\psi_i] = -\partial_i \phi, \quad [Q,N] = i\phi M, \quad [Q,\phi] = 0.
\end{align*}
\]

We assume that both $A$ and $M$ have ghost number 0, and thus will be regarded as bosonic fields when we study their quantum field theory. The ghost numbers of other fields can be read off the above transformation rules. We have that $\psi$ and $N$ are of ghost number 1, thus are fermionic, and $\phi$ is of ghost number 2 and bosonic. Note that the multiplet $(A, \psi, \phi)$ is what one would get in the topological field theory for Donaldson invariants except that our gauge group is $U(1)$, while the existence of $M$ and $N$ is a new feature. Also note that both $M$ and $\psi$ have the wrong statistics.

In order to construct a quantum field theory which will reproduce the SW invariants as correlation functions, anti–ghosts and Lagrangian multipliers are also required. We introduce the anti–ghost multiplet $(\lambda, \eta) \in \Omega^0(X)$, such that

\[
\begin{align*}
&[U,\lambda] = -2\lambda, \quad [Q,\lambda] = \eta, \quad [Q,\eta] = 0,
\end{align*}
\]

and the Lagrangian multipliers $(\chi, H) \in \Omega^{2+}(X)$, and $(\mu, \nu) \in S^- \otimes L$ such that

\[
\begin{align*}
&[U,\chi] = -\chi, \quad [Q,\chi] = H, \quad [Q,H] = 0; \\
&[U,\mu] = -\mu, \quad [Q,\mu] = \nu, \quad [Q,\nu] = i\phi \mu.
\end{align*}
\]

With the given fields, we construct the following functional which has ghost number -1:

\[
V = \int_X \left\{ \nabla_k \psi^k + \frac{i}{2} (\overline{N} M - M \overline{N}) ] \lambda - \chi^{kl} \left( H_{kl} - F^+_{kl} - \frac{i}{2} \overline{\lambda} \Gamma_{kl} M \right) \\
- \overline{\mu} (\nu - i D_{\lambda} M) - (\nu - i D_{\lambda} M) \mu \right\},
\]

(3.55)
where the indices of the tensorial fields are raised and lowered by a given metric \( g \) on \( X \), and the integration measure is the standard \( \sqrt{g}d^4x \). Also, \( \overline{M} \) and \( \bar{\mu} \) etc. represent the Hermitian conjugate of the spinorial fields. In a formal language, \( \overline{M} \in S^+ \otimes L^{-1} \) and \( \bar{\mu}, \bar{\nu} \) etc. represent the Hermitian conjugate of the spinorial fields. Following the standard procedure in constructing topological quantum field theory, we take the classical action of our theory to be [Zhang et al. (1995)]:

\[
S = \left\{ \phi, V \right\}, \\
\text{which has ghost number 0. One can easily show that } S \text{ is also BRST invariant, i.e., } [Q, S] = 0, \text{ thus it is invariant under the full super–algebra (3.54).}
\]

The bosonic Lagrangian multiplier fields \( H \) and \( \nu \) do not have any dynamics, and so can be eliminated from the action by using their equations of motion

\[
H_{kl} = \frac{1}{2} \left( \mathcal{F}_{kl}^+ + \frac{i}{2} \overline{M} \Gamma_{kl} M \right), \quad \nu = \frac{1}{2} i D_A M. \tag{3.56}
\]

Then we arrive at the following expression for the action [Zhang et al. (1995)]

\[
S = \int_X \left\{ \left[ -\Delta \phi + \overline{M} M \phi - i \overline{N} N \right] \lambda - \left[ \nabla_k \psi^k + \frac{i}{2} \left( \overline{N} M - \overline{M} N \right) \right] \eta \\
+ 2i \phi \bar{\mu} + (i D_A N - \gamma.\psi M) \bar{\mu} - \bar{\mu} (i D_A N - \gamma.\psi M) \right. \\
- \chi^{kl} \left[ \left( \nabla_k \psi^l - \nabla_l \psi^k \right) + \frac{i}{2} \left( \overline{M} \Gamma_{kl} N + \overline{N} \Gamma_{kl} M \right) \right] \left. \right\} + S_0,
\tag{3.57}
\]

where \( S_0 \) is given by

\[
S_0 = \int_X \left\{ \left\{ \frac{1}{4} |\mathcal{F}^+|^2 + \frac{i}{2} |\overline{M} \Gamma M|^2 + \frac{1}{2} |D_A M|^2 \right\} \right\}.
\]

It is interesting to observe that \( S_0 \) is nonnegative, and vanishes if and only if \( A \) and \( M \) satisfy the SW monopole equations. As pointed out in [Witten (1994)], \( S_0 \) can be rewritten as

\[
S_0 = \int_X \left\{ \left\{ \frac{1}{4} |\mathcal{F}^+|^2 + \frac{1}{4} |M|^4 + \frac{1}{8} R|M|^2 + g\mu^2 \mathcal{D}_A M \mathcal{D}_A M \right\} \right\},
\]

where \( R \) is the scalar curvature of \( X \) associated with the metric \( g \). If \( R \) is nonnegative over the entire \( X \), then the only square integrable solution of the monopole equations (3.52) is \( A \) a anti-self-dual connection and \( M = 0 \).
3.2.3.3 Quantum Field Theory

We will now investigate the quantum field theory defined by the classical action (3.57) with the path integral method. Let $\mathcal{F}$ collectively denote all the fields. The partition function of the theory is defined by [Zhang et al. (1995)]

$$Z = \int \mathcal{D}[\mathcal{F}] \exp \left( -\frac{1}{e^2} S \right),$$

where $e \in \mathbb{R}$ is the coupling constant. The integration measure $\mathcal{D}[\mathcal{F}]$ is defined on the space of all the fields. However, since $S$ is invariant under the gauge transformations, we assume the integration over the gauge field to be performed over the gauge orbits of $A$. In other words, we fix a gauge for the $A$ field using, say, a Faddeev–Popov procedure. This can be carried out in the standard manner, thus there is no need for us to spell out the details here. The integration measure $\mathcal{D}[\mathcal{F}]$ can be shown to be invariant under the super charge $Q$. Also, it does not explicitly involve the metric $g$ of $X$ [Zhang et al. (1995)].

Let $W$ be any operator in the theory. Its correlation function is defined by

$$Z[W] = \int \mathcal{D}[\mathcal{F}] \exp \left( -\frac{1}{e^2} S \right) W.$$

It follows from the $Q$ invariance of both the action $S$ and the path integration measure that for any operator $W$,

$$Z[[Q, W]] = \int \mathcal{D}[\mathcal{F}] \exp \left( -\frac{1}{e^2} S \right) [Q, W] = 0.$$

For the purpose of constructing topological invariants of the 4–manifold $X$, we are particularly interested in operators $W$ which are BRST–closed,

$$[Q, W] = 0,$$  \hspace{1cm} (3.58)

but not BRST–exact, i.e., can not be expressed as the (anti)–commutators of $Q$ with other operators. For such a $W$, if its variation with respect to the metric $g$ is BRST exact,

$$\delta_g W = [Q, W'],$$  \hspace{1cm} (3.59)
then its correlation function $Z[W]$ is a topological invariant of $X$ (by that we really mean that it does not depend on the metric $g$):

$$\delta_g Z[W] = \int \mathcal{D}[\mathcal{F}] \exp \left( -\frac{1}{e^2} S \right) [Q, W'] - \frac{1}{e^2} \delta_g V.W = 0.$$ 

In particular, the partition function $Z$ itself is a topological invariant.

Another important property of the partition function is that it does not depend on the coupling constant $e$:

$$\frac{\partial Z}{\partial e^2} = \int \mathcal{D}[\mathcal{F}] \left( \frac{1}{e^4} \exp \left( -\frac{1}{e^2} S \right) \right) [Q, V] = 0.$$ 

Therefore, $Z$ can be computed exactly in the limit when the coupling constant goes to zero. Such a computation can be carried out in the standard way: Let $A^\circ, M^\circ$ be a solution of the equations of motion of $A$ and $M$ arising from the action $S$. We expand the fields $A$ and $M$ around this classical configuration,

$$A = A^\circ + ea, \quad M = M^\circ + em,$$

where $a$ and $m$ are the quantum fluctuations of $A$ and $M$ respectively. All the other fields do not acquire background components, thus are purely quantum mechanical. We scale them by the coupling constant $e$, by setting $N$ to $eN$, $\phi$ to $e\phi$ etc.. To the order $o(1)$ in $e^2$, we have

$$Z = \sum_p \exp \left( -\frac{1}{e^2} S_{cl}^{(p)} \right) \int \mathcal{D}[\mathcal{F}'] \exp \left( -S_{q}^{(p)} \right),$$

where $S_{cl}^{(p)}$ is the quadratic part of the action in the quantum fields and depends on the gauge orbit of the classical configuration $A^\circ, M^\circ$, which we label by $p$. Explicitly, we have (see Zhang et al. (1995),

$$S_{cl}^{(p)} = \int_X \left\{ [-\Delta \phi + \bar{M}^o M^o \phi - i\bar{N}N] \lambda - \left[ \nabla_k \psi^k + \frac{1}{2} (\bar{N} M^o - \bar{M} N) \right] \eta \right. + 2i\phi \mu + (iD_{A^o} N - \gamma.\psi M^o) \mu - \bar{\mu} (iD_{A^o} N - \gamma.\psi M^o) \\
- \chi^{kl} \left[ \left( \nabla_k \psi^l - \nabla_l \psi^k \right)^+ + \frac{1}{2} (\bar{M}^o \Gamma_{kl} N + \bar{N} \Gamma_{kl} M^o) \right] \\
+ \frac{1}{4} \left| f^+ + i(\bar{m} \Gamma M^o + \bar{M}^o \Gamma m) \right|^2 + \frac{1}{2} \left| iD_{A^o} m + \gamma. a M^o \right|^2 \right\},$$

with $f^+$ the self-dual part of $f = da$. The classical part of the action is given by $S_{cl}^{(p)} = S_0|_{A=A^o, M=M^o}$. The integration measure $\mathcal{D}[\mathcal{F}']$ has exactly
the same form as $\mathcal{D}[\mathcal{F}]$ but with $A$ replaced by $a$, and $M$ by $m$, $\bar{M}$ by $\bar{m}$ respectively. Needless to say, the summation over $p$ runs through all gauge classes of classical configurations.

Let us now examine further features of our quantum field theory. A gauge class of classical configurations may give a non-zero contribution to the partition function in the limit $e^2 \to 0$ only if $S^{(p)}_q$ vanishes, and this happens if and only if $A^0$ and $M^0$ satisfy (3.52). Therefore, the SW monopole equations are recovered from the quantum field theory.

The equations of motion of the fields $\psi$ and $N$ in the semi-classical approximation can be easily derived from the quadratic action $S^{(p)}_q$, solutions of which are the zero modes of the quantum fields $\psi$ and $N$. The equations of motion read

\begin{align}
(d\psi)_{kl} + \frac{i}{2} \left( \bar{M}^o \Gamma_{kl} N + \bar{N} \Gamma_{kl} M^0 \right) &= 0, \\
D_{A^0} N + \gamma.\psi M^0 &= 0, \\
\nabla_k \psi^k + \frac{i}{2} (NM - \bar{M}N) &= 0.
\end{align}

Note that they are exactly the same equations which we have already discussed in (3.53). The first two equations are the linearization of the monopole equations, while the last is a ‘gauge fixing condition’ for $\psi$. The dimension of the space of solutions of these equations is the virtual dimension of the moduli space $\mathcal{M}$. Thus, within the context of our quantum field theoretical model, the virtual dimension of $\mathcal{M}$ is identified with the number of the zero modes of the quantum fields $\psi$ and $N$.

For simplicity we assume that there are no zero modes of $\psi$ and $N$, i.e., the moduli space is zero-dimensional. Then no zero modes exist for the other two fermionic fields $\chi$ and $\mu$. To compute the partition function in this case, we first observe that the quadratic action $S^{(p)}_q$ is invariant under the supersymmetry obtained by expanding $Q$ to first order in the quantum fields around the monopole solution $A^0$, $M^0$ (equations of motion for the nonpropagating fields $H$ and $\nu$ should also be used.). This supersymmetry transforms the set of 8 real bosonic fields (each complex field is counted as two real ones; the $a_i$ contribute 2 upon gauge fixing.) and the set of 16 fermionic fields to each other. Thus at a given monopole background we get [Zhang et al. (1995)]

\[ \int \mathcal{D}[\mathcal{F}]' \exp \left( -S^{(p)}_q \right) = \frac{\text{Pfaff}(\nabla \mathcal{F})}{|\text{Pfaff}(\nabla \mathcal{F})|} = \epsilon^{(p)}, \]

where $\epsilon^{(p)}$ is +1 or -1. In the above equation, $\nabla \mathcal{F}$ is the skew symmetric
first order differential operator defining the fermionic part of the action $S_q^{(p)}$, which can be read off from $S_q^{(p)}$ to be $\nabla_{\mathcal{F}} = \begin{pmatrix} 0 & T \\ -T^* & 0 \end{pmatrix}$. Therefore, $\epsilon^{(p)}$ is the sign of the determinant of the elliptic operator $T$ at the monopole background $A^m, M^a$, and the partition function $Z = \sum_p \epsilon^{(p)}$ coincides with the SW invariant of the 4–manifold $X$.

When the dimension of the moduli space $\mathcal{M}$ is greater than zero, the partition function $Z$ vanishes identically, due to integration over zero modes of the fermionic fields. In order to get any non trivial topological invariants for the underlying manifold $X$, we need to examine correlations functions of operators satisfying equations (3.58) and (3.59). A class of such operators can be constructed following the standard procedure [Zhang et al. (1995)]. We define the following set of operators

$$W_{k,0} = \frac{\phi^k}{k!}, \quad W_{k,1} = \psi W_{k-1,0},$$

$$W_{k,2} = \mathcal{F} W_{k-1,0} - \frac{1}{2} \psi \wedge \psi W_{k-2,0},$$

$$W_{k,3} = \mathcal{F} \wedge \psi W_{k-2,0} - \frac{1}{3!} \psi \wedge \psi \wedge \psi W_{k-3,0},$$

$$W_{k,4} = \frac{1}{2} \mathcal{F} \wedge \mathcal{F} W_{k-2,0} - \frac{1}{2} \mathcal{F} \wedge \psi \wedge \psi W_{k-3,0} - \frac{1}{4!} \psi \wedge \psi \wedge \psi \wedge \psi W_{k-4,0}.$$ (3.61)

These operators are clearly independent of the metric $g$ of $X$. Although they are not BRST invariant except for $W_{k,0}$, they obey the following equations [Zhang et al. (1995)]

$$dW_{k,0} = -[Q, W_{k,1}], \quad dW_{k,1} = [Q, W_{k,2}],$$

$$dW_{k,2} = -[Q, W_{k,3}], \quad dW_{k,3} = [Q, W_{k,4}], \quad dW_{k,4} = 0,$$

which allow us to construct BRST invariant operators from the the $W$’s in the following way: Let $X_i, i = 1, 2, 3, X_4 = X$, be compact manifolds without boundary embedded in $X$. We assume that these submanifolds are homologically nontrivial. Define

$$\hat{O}_{k,0} = W_{k,0}, \quad \hat{O}_{k,i} = \int_{X_i} W_{k,i}, \quad (i = 1, 2, 3, 4).$$ (3.62)

As we have already pointed out, $\hat{O}_{k,0}$ is BRST invariant. It follows from the descendent equations that

$$[Q, \hat{O}_{k,i}] = \int_{X_i} [Q, W_{k,i}] = \int_{X_i} dW_{k,i-1} = 0.$$
Therefore the operators $\hat{O}$ indeed have the properties (3.58) and (3.59). Also, for the boundary $\partial K$ of an $i + 1$D manifold $K$ embedded in $X$, we have

$$\int_{\partial K} W_{k,i} = \int_K dW_{k,i} = \left[ Q, \int_K W_{k,i+1} \right],$$

is BRST trivial. The correlation function of $\int_{\partial K} W_{k,i}$ with any BRST invariant operator is identically zero. This in particular shows that the $\hat{O}$'s only depend on the homological classes of the submanifolds $X_i$.

### 3.2.3.4 Dimensional Reduction and 3D Field Theory

In this subsection we dimensionally reduce the quantum field theoretical model for the SW invariant from 4D to 3D, thus to get a new topological quantum field theory defined on 3–manifolds. Its partition function yields a 3–manifold invariant, which can be regarded as the SW version of Casson’s invariant [Akbulut and McCarthy (1990); Taubes (1994)].

We take the 4–manifold $X$ to be of the form $Y \times [0, 1]$ with $Y$ being a compact 3–manifold without boundary. The metric on $X$ will be taken to be

$$(ds)^2 = (dt)^2 + g_{ij}(x)dx^i dx^j,$$

where the ‘time’ $t$–independent $g(x)$ is the Riemannian metric on $Y$. We assume that $Y$ admits a spin structure which is compatible with the $Spin_c$ structure of $X$, i.e., if we think of $Y$ as embedded in $X$, then this embedding induces maps from the $Spin_c$ bundles $S^\pm \otimes L$ of $X$ to $\tilde{S} \otimes L$, where $\tilde{S}$ is a spin bundle and $L$ is a line bundle over $Y$.

To perform the dimensional reduction, we impose the condition that all fields are $t$–independent. This leads to the following action [Zhang et al. (1995)]

$$S = \int \sqrt{g} d^4x \left\{ -\Delta \phi + \overline{M}M\phi - i\overline{N}N\lambda - [\nabla_k \psi^k + \frac{i}{2}(\overline{N}M - \overline{M}N)]\eta \
+ 2i\phi \mu \mu + [i(D_A + b)N - (\sigma.\psi - \tau)M]\mu - \overline{\mu} [i(D_A + b)N - (\sigma.\psi - \tau)M]\right. \
- 2\chi^k \left[ -\partial_k \tau + \ast(\nabla \psi)_k - \overline{M}\sigma_k N - \overline{N}\sigma_k M \right] \
+ \frac{1}{4} | \ast F - \partial b - M\sigma M |^2 + \frac{1}{2} | (D_A + b)M |^2 \right\},$$

where the $k$ is a 3D index, and $\sigma_k$ are the Pauli matrices. The fields $b, \tau \in \Omega^0(Y)$ respectively arose from $A_0$ and $\psi_0$ of the 4D theory, while the
meanings of the other fields are clear. The BRST symmetry in 4D carries over to the 3D theory. The BRST transformations rules for \((A_i, \psi_i, \phi)\), \(i = 1, 2, 3\), \((M, N)\), and \((\lambda, \eta)\) are the same as before, but for the other fields, we have

\[
[Q, b] = \tau, \quad [Q, \tau] = 0, \\
[Q, \chi_k] = \frac{1}{2} (\ast F_k - \partial_k b - M \sigma_k M), \\
[Q, \mu] = \frac{1}{2} i(D_A + b)M.
\]

The action \(S\) is cohomological in the sense that \(S = [Q, V_3]\), with \(V_3\) being the dimensionally reduced version of \(V\) defined by (3.55), and \([Q, S] = 0\). Thus it gives rise to a topological field theory upon quantization. The partition function of the theory

\[
Z = \int \mathcal{D}F \exp \left(-\frac{1}{e^2} S\right),
\]

can be computed exactly in the limit \(e^2 \to 0\), as it is coupling constant independent. We have, as before,

\[
Z = \sum_p \exp \left(-\frac{1}{e^2} S^{(p)}_{cl}\right) \int \mathcal{D}F' \exp \left(-S^{(p)}_q\right),
\]

where \(S^{(p)}_{cl}\) is the quadratic part of \(S\) expanded around a classical configuration with the classical parts for the fields \(A, M, b\) being \(A^o, M^o, b^o\), while those for all the other fields being zero. The classical action \(S^{(p)}_{cl}\) is given by

\[
S^{(p)}_{cl} = \int_Y \left\{ \frac{1}{4} |\ast F^o - \bar{M}^o \sigma M^o|^2 + \frac{1}{2} |D_{A^o} + b^o| M^o|^2 \right\},
\]

which can be rewritten as [Zhang et al. (1995)]

\[
S^{(p)}_{cl} = \int_Y \left\{ \frac{1}{4} |\ast F^o - \bar{M}^o \sigma M^o|^2 + \frac{1}{2} |D_{A^o} M^o|^2 + \frac{1}{2} |D_{b^o} + b^o|^2 + \frac{1}{2} |b^o M^o|^2 \right\}.
\]

In order for the classical configuration to have non-vanishing contributions to the partition function, all the terms in \(S^{(p)}_{cl}\) should vanish separately. Therefore,

\[
\ast F^o - \bar{M}^o \sigma M^o = 0, \quad D_{A^o} M^o = 0, \quad b^o = 0,
\]

(3.64)
where the last condition requires some explanation. When we have a trivial solution of the equations (3.64), it can be replaced by the less stringent condition $db = 0$. However, in a more rigorous treatment of the problem at hand, we in general perturb the equations (3.64), then the trivial solution does not arise.

Let us define an operator

$\tilde{T} : \Omega^0(Y) \oplus \Omega^1(Y) \oplus (\tilde{S} \otimes L) \to \Omega^0(Y) \oplus \Omega^1(Y) \oplus (\tilde{S} \otimes L)$,

$(\tau, \psi, N) \mapsto (-d^* \psi + \frac{i}{2}(\bar{N}M - \bar{M}N), \quad *d\psi - d\tau - \bar{N}\sigma M - \bar{M}\sigma N, \quad iD_A N - (\sigma, \psi - \tau)M)$,

(3.65)

where the complex bundle $\tilde{S} \otimes L$ should be regarded as a real one with twice the rank. This operator is self-adjoint, and is also obviously elliptic. We will assume that it is Fredholm as well. In terms of $\tilde{T}$, the equations of motion of the fields $\chi^i$ and $\mu$ can be expressed as $\tilde{T}^{(p)}(\tau, \psi, N) = 0$, where $\tilde{T}^{(p)}$ is the operator $\tilde{T}$ with the background fields $(A^o, M^o)$ belonging to the gauge class $p$ of classical configurations.

When the kernel of $\tilde{T}$ is zero, the partition function $Z$ does not vanish identically. An easy computation leads to $Z = \sum_p \epsilon^{(p)}$, where the sum is over all gauge inequivalent solutions of (3.64), and $\epsilon^{(p)}$ is the sign of the determinant of $\tilde{T}^{(p)}$.

A rigorous definition of the sign of the det($\tilde{T}$) can be devised. However, if we are to compute only the absolute value of $Z$, then it is sufficient to know the sign of det($\tilde{T}$) relative to a fixed gauge class of classical configurations. This can be achieved using the mod $-2$ spectral flow of a family of Fredholm operators $\tilde{T}_t$ along a path of solutions of (3.64). More explicitly, let $(A^o, M^o)$ belong to the gauge class of classical configurations $p$, and $(\tilde{A}^o, \tilde{M}^o)$ in $\tilde{p}$. We consider the solution of the SW equation on $X = Y \times [0, 1]$ with $A_0 = 0$ and also satisfying the following conditions

$$(A, M)|_{t=0} = (A^o, M^o), \quad (A, M)|_{t=1} = (\tilde{A}^o, \tilde{M}^o).$$

Using this solution in $\tilde{T}$ results in a family of Fredholm operators, which has zero kernels at $t = 0$ and 1. The spectral flow of $\tilde{T}_t$, denoted by $q(p, \tilde{p})$, is defined to be the number of eigenvalues which cross zero with a positive slope minus the number which cross zero with a negative slope. This number is a well defined quantity, and is given by the index of the operator $\frac{\partial}{\partial t} - \tilde{T}_t$ $[Zhang et al. (1995)]$. In terms of the spectral flow, we
have
\[ \det(\tilde{T}(p)) \det(\tilde{T}(\tilde{p})) = (-1)^{q(p,\tilde{p})}. \]

Equations (3.64) can be derived from the functional
\[ S_{c-s} = \frac{1}{2} \int_Y A \wedge F + i \int_Y \sqrt{g} d^3x M A. \]

Note that this is almost the standard Lagrangian of a $U(1)$ Chern–Simons theory coupled to spinors, except that we have taken $M$ to have bosonic statistics. $S_{c-s}$ is gauge invariant modulo a constant arising from the Chern–Simons term upon a gauge transformation. Therefore, \((\delta S_{c-s}, \delta A, \delta \bar{M})\) defines a vector field on the quotient space of all $U(1)$ connections $A$ tensored with the $\tilde{S} \times L$ sections by the $U(1)$ gauge group $G$, i.e., $W = (A \times \tilde{S} \times L)/G$. Solutions of (3.64) are zeros of this vector field, and \(\tilde{T}(p)\) is the Hessian at the point $p \in W$. Thus the partition $Z$ is nothing else but the Euler character of $W$. This geometrical interpretation will be spelt out more explicitly in the next subsection by re–interpreting the theory using the Mathai–Quillen formula [Mathai and Quillen (1986)].

3.2.3.5 Geometrical Interpretation

To elucidate the geometric meaning of the 3D theory obtained in the last section, we now cast it into the framework of Atiyah and Jeffrey [Atiyah and Jeffrey (1990)]. Let us briefly recall the geometric set up of the Mathai–Quillen formula as reformulated in [Atiyah and Jeffrey (1990)]. Let $P$ be a Riemannian manifold of dimension $2m + \dim G$, and $G$ be a compact Lie group acting on $P$ by isometries. Then $P \to P/G$ is a principle bundle. Let $V$ be a $2m$ dimensional real vector space, which furnishes a representation $G \to SO(2m)$. Form the associated vector bundle $P \times_G V$. Now the Thom form of $P \times_G V$ can be expressed [Zhang et al. (1995)]

\[ U = \exp(-x^2) \int \exp \{ i\chi \phi / 4 + i\chi dx - i\delta \nu, \lambda \} \] \[ \cdot \{ D\eta D\chi D\phi D\lambda, \] \[ (3.66) \]

where $x = (x^1, ..., x^{2m})$ is the coordinates of $V$, $\phi$ and $\lambda$ are bosonic variables in the Lie algebra $g$ of $G$, and $\eta$ and $\chi$ are Grassmannian variables valued in the Lie algebra and the tangent space of the fiber respectively. In the above equation, $C$ maps any $\eta \in g$ to the element of the vertical part.
Quantum Leap

of $\mathcal{T}P$ generated by $\eta$; $\nu$ is the $g$-valued 1–form on $P$ defined by $\langle \nu(\alpha), \eta \rangle = \langle \alpha, C(\eta) \rangle$, for all vector fields $\alpha$; and $R = C^*C$. Also, $\delta$ is the exterior derivative on $P$.

Now we choose a $G$ invariant map $s : P \to V$, and pull back the Thom form $U$. Then the top form on $P$ in $s^*U$ is the Euler class. If $\{\delta p\}$ forms a basis of the cotangent space of $P$ (note that $\nu$ and $\delta s$ are one forms on $P$), we replace it by a set of Grassmannian variables $\{\psi\}$ in $s^*U$, then intergrade them away. We arrive at

$$\Upsilon = \frac{1}{(2\pi)^{\dim G\frac{m}{2}}} \int \exp \left\{ -|s|^2 + i\chi \phi \chi/4 + i\chi \delta s - i\langle \delta \nu, \lambda \rangle - \langle \phi, R\lambda \rangle + \langle \psi, C\eta \rangle \right\} \text{D}\eta \text{D}\chi \text{D}\phi \text{D}\psi, \quad (3.67)$$

the precise relationship of which with the Euler character of $P \times_G V$ is

$$\int_P \Upsilon = \text{Vol}(G) \chi(P \times_G).$$

It is rather obvious that the action $S$ defined by (3.57) for the 4D theory can be interpreted as the exponent in the integrand of (3.67), if we identify $P$ with $\mathcal{A} \times \Gamma(W^+)$, and $V$ with $\Omega^{2,+}(X) \times \Gamma(W^-)$, and set $s = (\mathcal{F}^+ + \frac{i}{2}MTM, D_A M)$. Here $\mathcal{A}$ is the space of all $U(1)$ connections of $\text{det}(W^+)$, and $\Gamma(W^\pm)$ are the sections of $S^\pm \otimes L$ respectively.

For the 3D theory, we wish to show that the partition function yields the Euler number of $W$. However, the tangent bundle of $W$ cannot be regarded as an associated bundle with the principal bundle, for which for the formulae (3.66) or (3.67) can readily apply, some further work is required.

Let $P$ be the principal bundle over $P/G$, $V$, $V'$ be two orthogonal representations of $G$. Suppose there is an embedding from $P \times_G V'$ to $P \times_G V$ via a $G$–map $\gamma(p) : V' \to V$ for $p \in P$. Denote the resulting quotient bundle as $E$. In order to derive the Thom class for $E$, one needs to choose a section of $E$, or equivalently, a $G$–map $s : P \to V$ such that $s(p) \in (\text{Im}(\gamma(p)))^\bot$. Then the Euler class of $E$ can be expressed as $\pi_*\rho^*U$, where $U$ is the Thom class of $P \times_G V$, $\rho$ is a $G$–map: $P \times V' \to P \times V$ defined by

$$\rho(p, \tau) = (p, \gamma(p)\tau + s(p)),$$

and $\pi_*$ is the integration along the fiber for the projection $\pi : P \times V' \to$
P/G. Explicitly, we have [Zhang et al. (1995)]

\[
\pi^* \rho^*(U) = \int \exp \left\{ -|\gamma(p)\tau + s(p)|^2 + i\chi \phi \chi + i\delta(\gamma(p)\tau + s(p)) \right. \\
- i(\delta \nu, \lambda) - \langle \phi, R\lambda \rangle + \langle \nu, C\eta \rangle \} D\chi D\phi D\tau D\eta D\lambda.
\] (3.68)

Consider the exact sequence

\[ 0 \to \pi \to (A \times \Gamma(W)) \times_G \Omega^0(Y) \to j(A, M) \to \Omega^1(Y) \times \Gamma(W), \]

where \( j(A, M) : b \mapsto (-db, bM) \) (assuming that \( M \neq 0 \)). Then the tangent bundle of \( A \times \Gamma(W) \) can be regarded as the quotient bundle

\[ (A \times \Gamma(W)) \times_G (\Omega^1(Y) \times \Gamma(W))/\text{Im}(j). \]

We define a vector field on \( A \times \Gamma(W) \) by

\[ s(A, M) = (\star F_A - \bar{M} \sigma M, D_A M), \]

which lies in \( \text{Im}(j)^\perp \):

\[
\int_Y (\star F_A - \bar{M} \sigma M) \wedge (-db) + \int_Y \sqrt{g} d^3x (D_A M, bM) = 0, \quad (3.69)
\]

where we have used the short hand notation \( \langle M_1, M_2 \rangle = \frac{1}{2}(M_1 M_2 + M_2 M_1) \).

Formally applying the formula (3.68) to the present infinite-dimensional situation, we get the Euler class \( \pi^* \rho^*(U) \) for the tangent bundle \( T(A \times \Gamma(W)) \), where \( \rho \) is the \( G \)-invariant map \( \rho \) is defined by

\[
\rho : \Omega^0(Y) \to \Omega^1(Y) \times \Gamma(W), \quad \rho(b) = (-db + \star F_A - \bar{M} \sigma M, (D_A + b) M),
\]

\( \pi \) is the projection \( (A \times \Gamma(W)) \times_G \Omega^0(Y) \to A \times_G \Gamma(W) \), and \( \pi^* \) signifies the integration along the fiber. Also \( U \) is the Thom form of the bundle

\[ (A \times \Gamma(W)) \times_G (\Omega^1(Y) \times \Gamma(W)) \to A \times_G \Gamma(W). \]

To get a concrete feel about \( U \), we need to explain the geometry of this bundle. The metric on \( Y \) and the Hermitian metric \( \langle ., . \rangle \) on \( \Gamma(W) \) naturally define a connection. The Maurer–Cartan connection on \( A \to A/G \) is flat while the Hermitian connection on has the curvature \( i\phi \mu \wedge \bar{\mu} \). This gives the expression of term \( i(\chi, \mu) \phi(\chi, \mu) \) in \( (3.67) \) in our case.

In our infinite-dimensional setting, the map \( C \) is given by

\[
C : \Omega^0(Y) \to T(A, M)(A \times \Gamma(W)), \quad C(\eta) = (-d\eta, i\eta M),
\]
Quantum Leap

and its dual is given by

\[ C^* : \Omega^1(Y) \times \Gamma(W) \longrightarrow \Omega^0(Y), \quad C^*(\psi, N) = -d^* \psi + \langle N, iM \rangle. \]

The one form \( \langle \nu, \eta \rangle \) on \( A \times \Gamma(W) \) takes the value

\[ \langle (\psi, N), C\eta \rangle = \langle -d^* \psi, \eta \rangle + \langle N, \iota M \rangle \eta \]

on the vector field \((\psi, N)\). We also easily get \( R(\lambda) = -\Delta \lambda + \langle M, M \rangle \lambda \), where \( \Delta = d^* d \).

Combining all the information together, we arrive at the following formula,

\[ \exp \left\{ -\frac{1}{2} \rho^2 + i(\chi, \mu) \delta \rho + 2i\phi \mu \bar{\mu} \right\} \]

\[ + \langle \Delta \phi, \lambda \rangle - \phi \lambda \langle M, M \rangle + i\langle N, N \rangle \lambda \]

\[ + \langle \nu, \eta \rangle \} D\chi D\phi D\lambda D\eta Db. \]

Note that the 1–form \( i(\chi, \mu) \delta \rho \) on \( A \times \Gamma(W) \times \Omega^0(Y) \) contacted with the vector field \((\phi, N, b)\) leads to

\[ 2\chi^k [-\partial_k \tau + i(\nabla \psi)_k - \bar{M}\sigma_k N - \bar{N}\sigma_k M] + 2\langle \mu, [i(D_A + b)N - (\sigma, \psi - \tau)M] \rangle \]

and the relation \( (3.69) \) gives \( |\rho|^2 = |iF - M\sigma M|^2 + |db|^2 + |D_A M|^2 + b^2 |M|^2 \).

Finally we get the Euler character \cite{Zhang et al. (1995)}

\[ \pi_* \rho^* (U) = \int \exp(-S) D\chi D\phi D\lambda D\eta Db. \]

where \( S \) is the action \( (3.63) \) of the 3D theory defined on the manifold \( Y \).

Integrating \( (3.71) \) over \( A \times G \Gamma(W) \) leads to the Euler number

\[ \sum_{[(A, M)] : \chi(A, M) = 0} e^{A, M}, \]

which coincides with the partition function \( Z \) of our 3D theory.

### 3.2.4 TQFTs Associated with SW–Monopoles

Recall that TQFTs are often used to study topological nature of manifolds. In particular, 3D and 4D TQFTs are well developed. The most well–known 3D TQFT would be the Chern–Simons theory, whose partion
function gives *Ray–Singer torsion* of 3–manifolds and the other topological invariants can be obtained as gauge invariant observables i.e., *Wilson loops*. The correlation functions can be identified with knot or link invariants e.g., Jones polynomial or its generalizations. On the other hand, in 4D, a twisted $N = 2$ supersymmetric YM theory developed by Witten [Witten (1988b)] also has a nature of TQFT. This YM theory can be interpreted as *Donaldson theory* and the correlation functions are identified with *Donaldson polynomials*, which classify smooth structures of topological 4–manifolds. A new TQFT on 4–manifolds was discovered in SW studies of electric–magnetic duality of supersymmetric gauge theory. Seiberg and Witten [Seiberg and Witten (1994a); Seiberg and Witten (1994b)] studied the electric–magnetic duality of $N = 2$ supersymmetric $SU(2)$ YM gauge theory, by using a version of *Montonen–Olive duality* and obtained exact solutions. According to this result, the exact *low energy effective action* can be determined by a certain elliptic curve with a parameter $u = \langle \text{Tr}(\phi)^2 \rangle$, where $\phi$ is a complex scalar field in the adjoint representation of the gauge group, describing the quantum moduli space. For large $u$, the theory is weakly–coupled and semi–classical, but at $u = \pm \Lambda^2$ corresponding to strong coupling regime, where $\Lambda$ is the dynamically generated mass scale, the elliptic curve becomes singular and the situation of the theory changes drastically. At these singular points, magnetically charged particles become massless. Witten showed that at $u = \pm \Lambda^2$ the TQFT was related to the moduli problem of counting the solution of the (Abelian) ‘Seiberg–Witten monopole equations’ [Witten (1994)] and it gave a dual description for the $SU(2)$ Donaldson theory.

It turns out that in 3D a particular TQFT of *Bogomol'nyi monopoles* can be obtained from a dimensional reduction of Donaldson theory and the partition function of this theory gives the so–called *Casson invariant* of 3–manifolds [Atiyah and Jeffrey (1990)].

Ohta [Ohta (1998)] discussed TQFTs associated with the 3D version of both Abelian and non–Abelian SW–monopoles, by applying *Batalin–Vilkovisky quantization* procedure. In particular, Ohta constructed the topological actions, topological observables and BRST transformation rules.

In this subsection, mainly following [Ohta (1998)], we will discuss TQFTs associated with both Abelian and non–Abelian SW–monopoles. We will use the following notation.

Let $X$ be a compact orientable Spin 4–manifold without boundary and $g_{\mu\nu}$ be its Riemannian metric tensor (with $g = \det g_{\mu\nu}$). Here we use $x_\mu$ as the local coordinates on $X$. $\gamma_\mu$ are Dirac’s gamma matrices and
\[ \sigma_{\mu\nu} = \frac{\gamma_{\mu} \gamma_{\nu}}{2} \] with \( \{ \gamma_{\mu}, \gamma_{\nu} \} = g_{\mu\nu} \). \( M \) is a Weyl fermion and \( \overline{M} \) is a complex conjugate of \( M \). (We will suppress spinor indices.) The Lie algebra \( \mathfrak{g} \) is defined by \( [T^a, T^b] = i f_{abc} T^c \), where \( T^a \) is a generator normalized as \( \text{Tr}(T^a T^b) = \delta^{ab} \). The symbol \( f_{abc} \) is a structure constant of \( \mathfrak{g} \) and is antisymmetric in its indices. The Greek indices \( \mu, \nu, \alpha \) etc run from 0 to 3. The Roman indices \( a, b, c, \ldots \) are used for the Lie algebra indices running from 1 to \( \text{dim} \mathfrak{g} \), whereas \( i, j, k, \ldots \) are the indices for space coordinates. Space–time indices are raised and lowered with \( g_{\mu\nu} \). The repeated indices are assumed to be summed. \( \epsilon_{\mu\nu\rho\sigma} \) is an antisymmetric tensor with \( \epsilon_{0123} = 1 \).

We often use the abbreviation of roman indices as \( \theta^a = \theta^a T^a \) etc., in order to suppress the summation over Lie algebra indices.

**Brief Review of TQFT**

Firstly, we give a brief review of TQFT (compare with Witten’s TQFT presented in subsection 3.2.2 above).

Let \( \phi \) be any field content. For a local symmetry of \( \phi \), we can construct a nilpotent BRST–operator \( Q_B \) (\( Q_B^2 = 0 \)). The variation of any functional \( \mathcal{O} \) of \( \phi \) is denoted by \( \delta \mathcal{O} = \{ Q_B, \mathcal{O} \} \), where the bracket \( \{ *, * \} \) represents a graded commutator, that is, if \( \mathcal{O} \) is bosonic, the bracket means a commutator \( [*,*] \) and otherwise it is an anti–bracket. Now, we can give the definition of topological field theory, as given in [Birmingham et al. (1991)]:

A topological field theory consists of:

1. a collection of Grassmann graded fields \( \phi \) on an \( nD \) Riemannian manifold \( X \) with a metric \( g \),
2. a nilpotent Grassmann odd operator \( Q \),
3. physical states to be \( Q^- \)–cohomology classes,
4. an energy–momentum tensor \( T_{\alpha\beta} \) which is \( Q^- \)–exact for some functional \( V_{\alpha\beta} \) such as

\[
T_{\alpha\beta} = \{ Q, V_{\alpha\beta}(\phi, g) \}.
\]

In this definition, \( Q \) is often identified with \( Q_B \) and is in general independent of the metric. Now, recall that there are two broad types of TQFTs satisfying this definition and they are classified into Witten–type [Witten (1994)] or Schwarz–type [Schwarz (1978)].

For Witten–type TQFT, the quantum action \( S_\theta \) which comprises the classical action, ghost and gauge fixing terms, can be represented by \( S_\theta = \{ Q_B, V \} \), for some function \( V \) of metric and fields and BRST charge \( Q_B \).
Under the metric variation $\delta g$ of the partition function $Z$, it is easy to see that

$$\delta g Z = \int \mathcal{D}\phi \, e^{-S_\phi} \left( -\frac{1}{2} \int_X d^n x \sqrt{g} \delta g^{\alpha\beta} T_{\alpha\beta} \right)$$

$$= \int \mathcal{D}\phi \, e^{-S_\phi} \mathcal{Q} \equiv \langle \{Q, \chi \} \rangle = 0,$$

where $\chi = -\frac{1}{2} \int_X d^n x \sqrt{g} \delta g^{\alpha\beta} V_{\alpha\beta}$. 

The last equality in (3.72) follows from the BRST invariance of the vacuum and means that $Z$ is independent of the local structure of $X$, that is, $Z$ is a topological invariant of $X$.

In general, for Witten type theory, $Q_B$ can be constructed by an introduction of a topological shift with other local gauge symmetry [Ohta (1998)]. For example, in order to get the topological YM theory on four manifold $M^4$, we introduce the shift in the gauge transformation for the gauge field $A^a_\mu$ such as $\delta A^a_\mu = D_\mu \theta^a + \epsilon^a_\mu$, where $D_\mu$ is a covariant derivative, $\theta^a$ and $\epsilon^a_\mu$ are the (Lie algebra valued) usual gauge transformation parameter and topological shift parameter, respectively. In order to see the role of this shift, let us consider the first Pontryagin class on $M^4$ given by

$$S = \frac{1}{8} \int_{M^4} \epsilon^{\mu\nu\rho\sigma} F^a_{\mu\nu} F^a_{\rho\sigma} d^4 x,$$  

where $F^a_{\mu\nu}$ is a field strength of the gauge field. We can easily check the invariance of (3.73) under the action of $\delta$. In this sense, (3.73) has a larger symmetry than the usual YM gauge symmetry. Taking this into account, we can construct the topological YM gauge theory. We can also consider similar ‘topological’ shifts for matter fields.

In addition, in general, Witten type topological field theory can be obtained from the quantization of some Langevin equations. This approach has been used for the construction of several topological field theories, e.g., supersymmetric quantum mechanics, topological sigma models or Donaldson theory (see Birmingham et al. (1991)).

On the other hand, Schwarz-type TQFT [Schwarz (1978)] begins with any metric independent classical action $S_c$ as a starting point, but $S_c$ is assumed not to be a total derivative. Then the quantum action (up to gauge fixing term) can be written by

$$S_q = S_c + \{Q, V(\phi, g)\},$$

where

$$V(\phi, g) = \int_X d^n x \sqrt{g} \phi \theta.$$
for some function $V$. For this quantum action, we can easily check the topological nature of the partition function, but note that the energy–momentum tensor contributes only from the second term in (3.74). One of the differences between Witten type and Schwarz type theories can be seen in this point. Namely, the energy–momentum tensor of the classical action for Schwarz type theory vanishes because it is derived as a result of metric variation.

Finally, we comment on the local symmetry of Schwarz type theory. Let us consider the Chern–Simons theory as an example. The classical action,

$$ S_{CS} = \int_{M^3} d^3 x \left( A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right), \quad (3.75) $$

is a topological invariant, which gives the second Chern class of 3–manifold $M^3$. As is easy to find, $S_{CS}$ is not invariant under the topological gauge transformation, although it is YM gauge invariant. Therefore the quantization is proceeded by the standard BRST method. This is a general feature of Schwarz–type TQFT.

### 3.2.4.1 Dimensional Reduction

First, let us recall the SW monopole equations in 4D. We assume that $X$ has Spin structure. Then there exist rank two positive and negative spinor bundles $S^\pm$. For Abelian gauge theory, we introduce a complex line bundle $L$ and a connection $A_\mu$ on $L$. The Weyl spinor $M(\bar{M})$ is a section of $S^+ \otimes L$ ($S^- \otimes L^{-1}$), hence $M$ satisfies the positive chirality condition $\gamma^5 M = M$. If $X$ does not have Spin structure, we introduce $\text{Spin}^c$ structure and $\text{Spin}^c$ bundles $S^\pm \otimes L$, where $L^2$ is a line bundle. In this case, $M$ should be interpreted as a section of $S^+ \otimes L$. Below, we assume Spin structure.

Recall that the 4D Abelian SW monopole equations are the following set of differential equations

$$ F^{+}_{\mu\nu} + \frac{i}{2} \bar{M} \sigma_{\mu\nu} M = 0, \quad i \gamma^\mu D_\mu M = 0, \quad (3.76) $$

where $F^{+}_{\mu\nu}$ is the self–dual part of the $U(1)$ curvature tensor

$$ F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad F^{+}_{\mu\nu} = P^{+}_{\mu\nu} F^{\rho\sigma}, \quad (3.77) $$

while $P^{+}_{\mu\nu}$ is the self–dual projector defined by

$$ P^{+}_{\mu\nu} = \frac{1}{2} \left( \delta_{\mu\rho} \delta_{\nu\sigma} + \frac{\sqrt{g}}{2} \epsilon_{\mu\nu\rho\sigma} \right). $$
Note that the second term in the first equation of (3.76) is also self-dual. On the other hand, the second equation in (3.76) is a twisted Dirac equation, whose covariant derivative $D_\mu$ is given by

$$D_\mu = \partial_\mu + \omega_\mu - iA_\mu,$$

where $\omega_\mu = \frac{1}{4} \omega^{\alpha\beta}[\gamma_\alpha, \gamma_\beta]$ is the spin connection 1-form on $X$.

In order to perform a reduction to 3D, let us first assume that $X$ is a product manifold of the form $X = Y \times [0,1]$, where $Y$ is a 3D compact manifold which has Spin structure. We may identify $t \in [0,1]$ as a ‘time’ variable, or, we assume $t$ as the zero-th coordinate of $X$, whereas $x_i (i = 1, 2, 3)$ are the coordinates on (space manifold) $Y$. Then the metric is given by

$$ds^2 = dt^2 + g_{ij}dx^i dx^j.$$

The dimensional reduction is proceeded by assuming that all fields are independent of $t$. (Below, we suppress the volume factor $\sqrt{g}$ of $Y$ for simplicity.)

First, let us consider the Dirac equation. After the dimensional reduction, the Dirac equation will be

$$\gamma^i D_i M - i\gamma^0 A_0 M = 0.$$

As for the first monopole equation, using (3.77) we find that

$$F_{i0} + \frac{1}{2} \epsilon_{i0jk} F^{jk} = -i\overline{M} \sigma_{i0} M, \quad F_{ij} + \epsilon_{ijk0} F^{kj} = -i\overline{M} \sigma_{ij} M. \quad (3.78)$$

Since the above two equations are dual each other, the first one, for instance, can be reduced to the second one by a contraction with the totally anti-symmetric tensor. Thus, it is sufficient to consider one of them. Here, we take the first equation in (3.78).

After the dimensional reduction, (3.78) will be

$$\partial_i A_0 - \frac{1}{2} \epsilon_{ijk} F^{jk} = -i\overline{M} \sigma_{i0} M, \quad (3.79)$$

where we have set $\epsilon_{ijk} \equiv \epsilon_{0ijk}$.

Therefore, the 3D version of the SW equations are given by

$$\partial_i b - \frac{1}{2} \epsilon_{ijk} F^{jk} + i\overline{M} \sigma_{i0} M = 0, \quad i(\gamma^i D_i - i\gamma^0 b)M = 0, \quad (b \equiv A_0). \quad (3.80)$$
Quantum Leap

It is now easy to establish the non–Abelian 3D monopole equations as

\[ \partial_i b^a + f_{abc} A^b_i b^c - \frac{1}{2} \epsilon_{ijk} F^{ajk} + i \overline{M} \sigma_{0a} M = 0, \quad i(\gamma^i D_i - i\gamma^0 b)M = 0, \]

where we have abbreviated \( \overline{M} \sigma_{\mu\nu} T^a M \equiv \overline{M} \sigma_{\mu\nu} (T^a)_{ij} M^j \), subscripts of \((T^a)_{ij}\) run from 1 to \( \text{dim} \ g \) and \( b^a \equiv A^a_0 \).

Next, let us find an action which produces (3.80). The simplest one is given by

\[ S = \frac{1}{2} \int_Y \left[ \left( \partial_i b - \frac{1}{2} \epsilon_{ijk} F^{jk} + i \overline{M} \sigma_{0a} M \right)^2 + |i(\gamma^i D_i - i\gamma^0 b)M|^2 \right] d^3x. \]  

(3.81)

Note that the minimum of (3.81) is given by (3.80). In this sense, the 3D monopole equations are not equations of motion but rather of constraints.

Furthermore, there is a constraint for \( b \). To see this, let us rewrite (3.81) as

\[ S = \int_Y d^3x \left[ \frac{1}{2} \left( \frac{1}{2} \epsilon_{ijk} F^{jk} - i \overline{M} \sigma_{0a} M \right)^2 + \frac{1}{4} |\gamma^i D_i M|^2 + \frac{1}{2} |\partial_i b|^2 + \frac{1}{2} b^2 |M|^2 \right]. \]

The minimum of this action is clearly given by the 3D monopole equations with \( b = 0 \), for non–trivial \( A_i \) and \( M \). However, for trivial \( A_i \) and \( M \), we may relax the condition \( b = 0 \) to \( \partial_i b = 0 \), i.e., \( b \) is in general a non–zero constant. This can be also seen from (3.79). Therefore, we get

\[ \frac{1}{2} \epsilon_{ijk} F^{jk} - i \overline{M} \sigma_{0a} M = 0, \quad i\gamma^i D_i M = 0, \]

with \( b = 0 \) or \( \partial_i b = 0 \),

(3.82)

as an equivalent expression to (3.80), but we will rather use (3.80) for convenience. The Gaussian action will be used in the next subsection in order to construct a TQFT by Batalin–Vilkovisky quantization algorithm. The non–Abelian version of (3.81) and (3.82) would be obvious.

3.2.4.2 TQFTs of 3D Monopoles

In this subsection, we construct TQFTs associated with both the Abelian and non–Abelian 3D monopoles by Batalin–Vilkovisky quantization algorithm.
Abelian Case

A 3D action for the Abelian 3D monopoles was found by the direct dimensional reduction of the 4D one [Zhang et al. (1995)], but here we rather show that the 3D topological action can be directly constructed from the 3D monopole equations [Ohta (1998)].

Topological Bogomol’nyi Action

A topological Bogomol’nyi action was constructed by using Batalin–Vilkovisky quantization algorithm [Birmingham et al. (1989)], or quantization of a magnetic charge [Baulieu and Grossman (1988)]. The former is based on the quantization of a certain Langevin equation (‘Bogomol’nyi monopole equation’) and the classical action is quadratic, but the latter is based on the ‘quantization’ of the pure topological invariant by using the Bogomol’nyi monopole equation as a gauge fixing condition.

In order to compare the action to be constructed with those of Bogomol’nyi monopoles [Birmingham et al. (1989); Baulieu and Grossman (1988)], we take Batalin–Vilkovisky procedure (see also [Birmingham et al. (1991)]).

In order to get the topological action associated with 3D monopoles, we introduce random Gaussian fields \( G_i \) and \( \nu(\bar{\nu}) \) and then start with the action

\[
S_c = \frac{1}{2} \int \left[ \left( G_i - \partial_i b + \frac{1}{2} \epsilon_{ijk} F_{jk} - i M \sigma_{0i} \right)^2 + \left( \nu - i \gamma^i D_i M - \gamma^0 b M \right)^2 \right] d^3 x. \tag{3.83}
\]

Note that \( G_i \) and \( \nu(\bar{\nu}) \) are also regarded as auxiliary fields. This action reduces to (3.81) in the gauge

\[ G_i = 0, \quad \nu = 0. \tag{3.84} \]

Firstly, note that (3.83) is invariant under the topological gauge transformation

\begin{align*}
\delta A_i &= \partial_i \theta + \epsilon_i, \quad \delta b = \tau, \quad \delta M = i \theta M + \varphi, \\
\delta G_i &= \partial_i \tau - \epsilon_{ijk} \partial_j \epsilon^k + i (\bar{\sigma} \sigma_{0i} M + M \sigma_{0i} \varphi), \\
\delta \nu &= i \theta \nu + \gamma^i \epsilon_i M + i \gamma^i D_i \varphi + \gamma^0 b \varphi + \gamma^0 M, \tag{3.85}
\end{align*}

where \( \theta \) is the parameter of gauge transformation, \( \epsilon_i \) and \( \tau \equiv \epsilon_4 \) are parameters which represent the topological shifts and \( \varphi \) the shift on the
spinor space. The brackets for indices means anti–symmetrization, i.e.,
\[ A_i[B_j] = A_iB_j - A_jB_i. \]

Here, let us classify the gauge algebra (3.85). This is necessary to use Batalin–Vilkovisky algorithm. Let us recall that the local symmetry for fields \( \phi_i \) can be written generally in the form
\[ \delta \phi_i = R_i^{\alpha}(\phi)e^\alpha, \]
where the indices mean the label of fields and \( e^\alpha \) is a some local parameter. When \( \delta \phi_i = 0 \) for non–zero \( e^\alpha \), this symmetry is called \textit{first–stage reducible}. In the reducible theory, we can find zero–eigenvectors \( Z_\alpha \) satisfying
\[ R_i^{\alpha}Z_\alpha = 0. \]
Moreover, when the theory is \textit{on–shell reducible}, we can find such eigenvectors by using equations of motion.

For the case at hand, under the identifications
\[ \theta = \Lambda, \quad \epsilon_i = -\partial_i \Lambda, \quad \varphi = -i\Lambda M, \quad \tau = 0, \]
so that (3.85) will be
\[ \delta A_i = 0, \quad \delta b = 0, \quad \delta M = 0, \quad \delta G_i = 0, \]
\[ \delta \nu = i\Lambda(\nu - i\gamma^4 D_i M - \gamma^0 b M) \big|_{\text{on–shell}} = 0. \]
Then for \( \delta A_i \), for example, the \( R \) coefficients and the zero–eigenvectors are derived from
\[ \delta A_i = R_i^A Z_\Lambda Z_\alpha^A + R_{ij}^A Z_{ij}^A = 0, \]
that is
\[ R_i^A = \partial_i, \quad R_{ij}^A = \delta_{ij}, \quad Z_\Lambda^A = 0, \quad Z_{ij}^A = -\partial_j. \]
Obviously, similar relations hold for other fields. The reader may think that the choice (6.93) is not suitable as a first stage reducible theory, but note that the zero–eigenvectors appear on every point where the gauge equivalence and the topological shift happen to coincide. In this three dimensional theory, \( b(A_0) \) is invariant for the usual infinitesimal gauge transformation because of its “time” independence, so (6.93) means that the existence of the points on spinor space where the topological shift trivializes indicates the first stage reducibility.

If we carry out \textit{BRST quantization via Faddeev–Popov procedure} in this situation, the Faddeev–Popov determinant will have zero modes. Therefore in order to fix the gauge further we need a ghost for ghost. This reflects on the second generation gauge invariance (3.87) realized on–shell. However, since \( b \) is irrelevant to \( \Lambda \), the ghost for \( \tau \) will not couple to the second generation ghost. With this in mind, we use Batalin–Vilkovisky algorithm in order to make BRST quantization (for details, see [Ohta (1998)] and references therein).
Let us assign new ghosts carrying opposite statistics to the local parameters. The assortment is given by
\[ \theta \rightarrow c, \quad \epsilon_i \rightarrow \psi_i, \quad \tau \rightarrow \xi, \quad \varphi \rightarrow N, \quad \Lambda \rightarrow \phi, \]  
(3.88)

Ghosts in (3.88) are first generations, in particular, \( c \) is Faddeev–Popov ghost, whereas \( \phi \) is a second generation ghost. Their Grassmann parity and ghost number (U number) are given by
\[ c \quad \psi_i \quad \xi \quad N \quad \phi \]
\[ 1^- \quad 1^- \quad 1^- \quad 1^- \quad 2^+, \]  
(3.89)

where the superscript of ghost number denotes the Grassmann parity. Note that the ghost number counts the degree of differential form on the moduli space \( M \) of the solution to the 3D monopole equations. The minimal set \( \Phi_{\text{min}} \) of fields consists of
\[ A_i \quad b \quad M \quad G_i \quad \nu \quad 0^+ \quad 0^+ \quad 0^+ \quad 0^+, \]  
(3.89)

On the other hand, the set of anti–fields \( \Phi^*_{\text{min}} \) carrying opposite statistics to \( \Phi_{\text{min}} \) is given by
\[ A_i^* \quad b^* \quad M^* \quad G_i^* \quad \nu^* \quad c^* \quad \psi_i^* \quad N^* \quad \phi^* \]
\[ -1^- \quad -1^- \quad -1^- \quad -1^- \quad -2^+ \quad -2^+ \quad -2^+ \quad -3^- \]  
(3.90)

Next step is to find a solution to the master equation with \( \Phi_{\text{min}} \) and \( \Phi^*_{\text{min}} \), given by
\[ \frac{\partial S}{\partial \Phi_A} \frac{\partial S}{\partial \Phi_A^*} - \frac{\partial S}{\partial \Phi_A^*} \frac{\partial S}{\partial \Phi_A} = 0, \]  
(3.90)

where \( r(l) \) denotes right (left) derivative.

The general solution for the first stage reducible theory at hand can be expressed by
\[ S = S_c + \Phi^*_i R_i A^\alpha C_1^\alpha + C^*_1 A^{\alpha \beta} C_2^\beta + T_1^{\beta \gamma} C_1^\gamma C_1^\beta + C_1^\alpha A^\gamma A_1^\beta C_1^{\alpha \beta} + \Phi^*_i \Phi^*_j B_i^\alpha C_2^\beta + \cdots, \]  
(3.91)

where \( C_1^\alpha (C_2^\beta) \) denotes generally the first (second) generation ghost and only relevant terms in our case are shown. We often use \( \Phi^A_{\text{min}} = (\phi^i, C_1^\alpha, C_2^\beta) \), where \( \phi^i \) denote generally the fields. In this expression, the indices should be interpreted as the label of fields. Do not confuse with
space–time indices. The coefficients $Z_\alpha^\beta, T_\alpha^\beta$, etc can be directly determined from the master equation. In fact, it is known that these coefficients satisfy the following relations

$$
\begin{align*}
R_i^\alpha Z_\alpha^\beta C_2^\beta - 2 \frac{\partial_i S_c}{\partial \phi^i} B_j^\alpha C_2^\alpha (-1)^{|i|} &= 0, \\
\frac{\partial_i Z_\alpha^\beta C_2^\beta}{\partial \phi^i} - 2 T_\alpha^\beta C_1^\beta + R_i^\alpha T_\alpha^\beta C_1^\gamma C_1^\delta &= 0, \\
\frac{\partial_i Z_\alpha^\beta C_2^\beta}{\partial \phi^i} - 2 T_\alpha^\beta C_1^\beta + 2 R_i^\alpha T_\alpha^\beta C_1^\gamma C_2^\delta + Z_\alpha^\beta A_\delta^\alpha C_1^\gamma C_1^\delta &= 0,
\end{align*}
$$

(3.92)

where $|i|$ means the Grassmann parity of the $i$th field.

In these expansion coefficients, $R_i^\alpha$ and $Z_\alpha^\beta$ are related to the local symmetry (3.85). On the other hand, as $T_\alpha^\beta$ is related to the structure constant of a given Lie algebra for a gauge theory, it is generally called as structure function. Of course if the theory is Abelian, such structure function does not appear. However, for a theory coupled with matters, all of the structure functions do not always vanish, even if the gauge group is Abelian. At first sight, this seems to be strange, but the expansion (3.91) obviously detects the coupling of matter fields and ghosts. In fact, the appearance of this type of structure function is required in order to make the action to be constructed being full BRST invariant.

After some algebraic works, we will find the solution to be

$$
S(\Phi_{\min}, \Phi^*_{\min}) = S_c + \int_Y \Delta S d^3x,
$$

where

$$
\Delta S = A_i^*(\partial^i c + \psi^i) + b^* \xi + M^*[i c M + N] + \overline{M}^*(-i c \overline{M} + \overline{N})
$$

$$
+ G_i^*[\partial^i \xi - \epsilon^{ijk} \partial_j \psi_k + i(\overline{N} \sigma^{0i} M + \overline{M} \sigma^{0i} N)]
$$

$$
+ \nu^*(i c \nu + i \gamma^i D_i N + \gamma^i \psi_i M + \gamma^0 b N + \gamma^0 \xi M)
$$

$$
+ \nu^*[i c \nu + i \gamma^i D_i N + \gamma^i \psi_i M + \gamma^0 b N + \gamma^0 \xi M]
$$

$$
+ c^* \phi - \psi_i^* \partial^i \phi - i N^*(\phi M + c N) + i \overline{N}^*(\phi \overline{M} + \overline{c} N) + 2 i \nu^* \nu^* \phi.
$$

We augment $\Phi_{\min}$ by new fields $\chi_i, d_i, \mu(\pi), \zeta(\overline{\zeta}), \lambda, \rho, \eta, \epsilon$ and the corresponding anti–fields. Their ghost number and Grassmann patity are given by

$$
\begin{array}{cccccccccc}
\chi_i & d_i & \mu & \zeta & \lambda & \rho & \eta & \epsilon \\
-1^- & 0^+ & -1^- & 0^+ & -2^+ & -1^- & -1^- & 0^+
\end{array}
$$

and

$$
\begin{array}{cccc}
\chi_i^* & \mu^* & \lambda^* & \rho^*
\end{array}
$$

0^+ & 1^- & 0^+.
Then we look for the solution

$$ S' = S(\Phi_{\text{min}}, \Phi^*_{\text{min}}) + \int_Y (\chi^* d_i + \pi^* \zeta + \mu^* \xi + \rho^* e + \lambda^* \eta) d^3 x, \quad (3.93) $$

where $d_i, \zeta, e, \eta$ are Lagrange multiplier fields.

In order to get the quantum action we must fix the gauge. The best choice for the gauge fixing condition, which can reproduce the action obtained from the dimensional reduction of the 4D one, is found to be [Ohta (1998)]

$$ G_i = 0, \quad \nu = 0, \quad \partial^i A_i = 0, \quad \rho \partial^i A_i = 0. $$

Thus we can get the gauge fermion carrying the ghost number $-1$ and odd Grassmann parity,

$$ \Psi = -\chi^* G_i - \pi^* \nu - \mu^* \pi - \rho \partial^i A_i - \lambda \left[ -\partial^i \psi_i + \frac{i}{2}(\mathcal{N} M - \mathcal{M} N) \right]. $$

The quantum action $S_q$ can be obtained by eliminating anti–fields and are restricted to lie on the gauge surface $\Phi^* = \partial_r \Psi$. Therefore the anti–fields will be

$$ G_i^* = -\chi_i, \quad \chi_i = -G_i, \quad \nu^* = -\pi, \quad \pi^* = -\mu, \quad \mu^* = -\nu, $$

$$ \rho^* = \pi, \quad M^* = -\frac{i}{2} \lambda N, \quad N^* = \frac{i}{2} \lambda M, $$

$$ \lambda^* = -\left[ -\partial^i \psi_i + \frac{i}{2}(\mathcal{N} M - \mathcal{M} N) \right], \quad e^* = \phi^* = b^* = \zeta^* (\zeta^*) = 0. \quad (3.94) $$

Then the quantum action $S_q$ is given by $S_q = S' (\Phi, \Phi^* = \partial_r \Psi / \partial \Phi)$. Substituting [3.94] into $S_q$, we find that

$$ S_q = S_c + \int_Y \tilde{\Delta} S d^3 x, \quad \text{where} $$
Quantum Leap

\[ \Delta S = (-\Delta \phi + \phi MM - iNN)\lambda - \left[ -\partial^I \phi_i + i \left( \frac{N}{2} M - \frac{M}{2} N \right) \right] \eta \]

\[ = -\bar{\epsilon} (\gamma^I, D_iN + \gamma^I \psi_i M + \gamma^I b N + \gamma^I \xi M) \]

\[ + \frac{N}{i} \left( \gamma^I D_iN + \gamma^I \psi_i M + \gamma^I b N + \gamma^I \xi M \right) \mu + 2i \phi \bar{\mu} \]

\[ - \chi^I \left[ \partial_i \xi - \epsilon_{ijk} \partial^j \psi_k + i(\bar{N} M \sigma \sigma_0 M + \bar{M} \sigma \sigma_0 N) \right] + \rho (\bar{\Delta} \nu + \partial^I \psi_i) \]

Using the condition \( (3.84) \) with \( c = 0 \), we can arrive at

\[ S'_q = S_{\mid \gamma = 0} + \int Y \Delta S \mid \gamma = 0 \, d^3 x, \quad \text{where} \quad \Delta S \mid \gamma = 0 \]

\[ = (-\Delta \phi + \phi MM - iNN)\lambda - \left[ -\partial^I \phi_i + i \left( \frac{N}{2} M - \frac{M}{2} N \right) \right] \eta \]

\[ = -\bar{\epsilon} (\gamma^I, D_iN + \gamma^I \psi_i M + \gamma^I b N + \gamma^I \xi M) \]

\[ + \frac{N}{i} \left( \gamma^I D_iN + \gamma^I \psi_i M + \gamma^I b N + \gamma^I \xi M \right) \mu + 2i \phi \bar{\mu} \]

\[ - \chi^I \left[ \partial_i \xi - \epsilon_{ijk} \partial^j \psi_k + i(\bar{N} M \sigma \sigma_0 M + \bar{M} \sigma \sigma_0 N) \right] + \rho (\bar{\Delta} \nu + \partial^I \psi_i) \]

It is easy to find that \( (3.96) \) is consistent with the action found by the dimensional reduction of the 4D topological action [Zhang et al. (1995)].

**BRST Transformation**

The Batalin–Vilkovisky algorithm also facilitates to construct BRST transformation rule. The BRST transformation rule for a field \( \Phi \) is defined by

\[ \delta_B \Phi = \epsilon \left. \frac{\partial S'}{\partial \Phi^*} \right|_{\Phi^* = \Phi^* - \frac{\partial S'}{\partial \Phi}}, \]

where \( \epsilon \) is a constant Grassmann odd parameter. With this definition for \( (4.104) \), we get

\[ \delta_B A_i = -\epsilon (\partial_i c + \psi_i), \quad \delta_B b = -\epsilon \xi, \quad \delta_B M = -\epsilon (i c M + N), \]

\[ \delta_B G_i = -\epsilon \left( \partial_i\xi - \epsilon_{ijk} \partial^j \psi_k + i(\bar{N} M \sigma \sigma_0 M + \bar{M} \sigma \sigma_0 N) \right), \]

\[ \delta_B \nu = -\epsilon (i c \nu + i \gamma^I D_i N + \gamma^I \psi_i M + \gamma^I b N + \gamma^I \xi M - i \mu \phi), \]

\[ \delta_B c = \epsilon \phi, \quad \delta_B \psi_i = -\epsilon \partial_i \phi, \quad \delta_B \mu = -\epsilon \eta, \]

\[ \delta_B \lambda = -\epsilon \lambda, \quad \delta_B N = -\epsilon (\phi M + c N), \quad \delta_B X_i = \epsilon d_i, \]

\[ \delta_B \phi = \delta_B \xi = \delta_B d_i = \delta_B c = \delta_B \zeta = \delta_B \eta = 0. \]
It is clear at this stage that (3.98) has on-shell nilpotency, i.e., the quantum equation of motion for $\nu$ must be used in order to have $\delta B = 0$. This is due to the fact that the gauge algebra has on–shell reducibility. Accordingly, the Batalin–Vilkovisky algorithm gives a BRST invariant action and on–shell nilpotent BRST transformation. Note that the equations

$$\partial_i \xi - \epsilon_{ijk} \partial^j \psi^k + i(\mathbf{N}\sigma_{a0}M + \mathbf{M}\sigma_{a0}N) = 0,$$

$$i\gamma^i D_i N + \gamma^i \psi_i M + \gamma^0 b N + \gamma^0 \xi M = 0$$

can be recognized as linearizations of the 3D monopole equations and the number of linearly independent solutions gives the dimension of the moduli space $\mathcal{M}$.

It is now easy to show that the global supersymmetry can be recovered from (3.98). In Witten type theory, $Q_B$ can be interpreted as a supersymmetric BRST charge. We define the supersymmetry transformation as

$$\delta_S \Phi := \delta_B \Phi |_{c=0}.$$  

**Off–Shell Action**

As was mentioned before, the quantum action of Witten type TQFT can be represented by BRST commutator with nilpotent BRST charge $Q_B$. However, since our BRST transformation rule is on-shell nilpotent, we should integrate out $\nu$ and $G_i$ in order to get off–shell BRST transformation and off–shell quantum action.

For this purpose, let us consider the following terms in (4.104),

$$\frac{1}{2}(G_i - X_i)^2 + \frac{1}{2}|\nu - A|^2 - i\pi c\nu + i\pi b - \bar{\nu} c\nu - \bar{\nu} - d^i G_i,$$

(3.99)

where $X_i = \partial_i b - \frac{1}{2} \epsilon_{ijk} F_{jk} + i\mathbf{M}\sigma_{a0}M$, $A = i\gamma^i D_i M + \gamma^0 b M$.

Here, let us define

$$\nu' = \nu - A, \quad B = -i c\mu - \zeta.$$  

$\nu'(\nu')$ and $G_i$ can be integrated out and then (3.99) will be

$$-\frac{1}{2} d_i d^i - d_i X_i - 2|B|^2 + \bar{B}A + B\bar{A}.$$  

Consequently, we get the off–shell quantum action

$$S_q = \{Q, \bar{\Psi}\},$$

where (3.100)
Quantum Leap

\[ \Psi = -\chi^i \left( X_i + \frac{\alpha}{2} A_i \right) - \mu (\gamma^i D_i M + \gamma^0 b M - \beta B) - \bar{\rho} (\gamma^i D_i M + \gamma^0 b M - \beta B) + \rho \partial^i A_i - \lambda \left[ -\partial^i \psi_i + \frac{i}{2} (NM - MN) \right]. \]

\(\alpha\) and \(\beta\) are arbitrary gauge fixing parameters. Convenience choice for them is \(\alpha = \beta = 1\). The BRST transformation rule for \(X_i\) and \(B\) fields can be easily obtained, although we do not write down here.

Observables

We can now discuss the observables. For this purpose, let us define [Baulieu and Grossman (1988)]

\[ A = A + c, \quad F = F + \psi - \phi, \quad K = db + \xi, \]

where we have introduced differential form notations, but their meanings would be obvious. \(A\) and \(c\) are considered as a \((1, 0)\) and \((0, 1)\) part of 1–form on \((Y, \mathcal{M})\). Similarly, \(F, \psi\) and \(\phi\) are \((2, 0)\), \((1, 1)\) and \((0, 2)\) part of the 2–form \(F\), and \(db\) and \(\xi\) are \((1, 0)\) and \((0, 1)\) part of the 1–form \(K\). Thus \(A\) defines a connection 1–form on \((Y, \mathcal{M})\) and \(F\) is a curvature 2–form. Note that the exterior derivative \(d\) maps any \((p_1, p_2)\)–form \(X_p\) of total degree \(p = p_1 + p_2\) to \((p_1 + 1, p_2)\)–form, but \(\delta B\) maps any \((p_1, p_2)\)–form to \((p_1, p_2 + 1)\)–form. Also note that \(X_p X_q = (-1)^{pq} X_q X_p\). Then the action of \(\delta B\) is

\[ (d + \delta B) A = F, \quad (d + \delta B) b = K. \quad (3.101) \]

\(F\) and \(K\) also satisfy the following Bianchi identities in Abelian theory:

\[ (d + \delta B) F = 0, \quad (d + \delta B) K = 0. \quad (3.102) \]

Equations \((3.101)\) and \((3.102)\) mean anti–commuting property between the BRST variation \(\delta B\) and the exterior differential \(d\), i.e., \(\{\delta B, d\} = 0\).

The BRST transformation rule in geometric sector can be easily read from \((3.98)\), i.e., \(\delta B A, \delta B \psi, \delta B c\) and \(\delta B \phi\). \((3.102)\) implies

\[ (d + \delta B) F^n = 0, \quad (3.103) \]

and expanding the above expression by ghost number and form degree, we
get the following \((i, 2n - i)\)-form \(W_{n,i}\),

\[
W_{n,0} = \frac{\phi^n}{n!}, \quad W_{n,1} = \frac{\phi^{n-1}}{(n-1)!} \psi, \quad W_{n,2} = \frac{\phi^{n-2}}{2(n-2)!} \psi \wedge \psi - \frac{\phi^{n-1}}{(n-1)!} F, \\
W_{n,3} = \frac{\phi^{n-3}}{6(n-3)!} \psi \wedge \psi \wedge \psi - \frac{\phi^{n-2}}{(n-2)!} F \wedge \psi, 
\]

(3.104)

where \(0 = \delta_B W_{n,0}\), \(dW_{n,0} = \delta_B W_{n,1}\), \(dW_{n,1} = \delta_B W_{n,2}\), \(dW_{n,2} = \delta_B W_{n,3}\), \(dW_{n,3} = 0\).

Picking a certain \(k\)-cycle \(\gamma\) as a representative and defining the integral

\[
W_{n,k}(\gamma) = \int_{\gamma} W_{n,k},
\]

we can easily prove that

\[
\delta_B W_{n,k}(\gamma) = -\int_{\gamma} dW_{n,k-1} = -\int_{\partial\gamma} W_{n,k-1} = 0,
\]

as a consequence of (3.105). Note that the last equality follows from the fact that the cycle \(\gamma\) is a simplex without boundary, i.e., \(\partial\gamma = 0\). Therefore, \(W_{n,k}(\gamma)\) indeed gives a topological invariant associated with \(n\)-th Chern class on \(Y \times M\).

On the other hand, since we have a scalar field \(b\) and its ghosts, we may construct topological observables associated with them. Therefore, the observables can be obtained from the ghost expansion of

\[
(d + \delta_B) F^n \wedge K^m = 0.
\]

Explicitly, for \(m = 1\), for example, we get

\[
0 = \delta_B W_{n,1,0}, \quad dW_{n,1,0} = \delta_B W_{n,1,1}, \quad dW_{n,1,1} = \delta_B W_{n,1,2}, \\
dW_{n,1,2} = \delta_B W_{n,1,3}, \quad dW_{n,1,3} = 0,
\]

where
Quantum Leap

\[ W_{n,1,0} = \frac{\phi^n}{n!} \xi, \quad W_{n,1,1} = \frac{\phi^{n-1}}{(n-1)!} \psi \xi - \frac{\phi^n}{n!} db, \]
\[ W_{n,1,2} = \frac{\phi^{n-2}}{2(n-2)!} \psi \wedge \psi \xi - \frac{\phi^{n-1}}{(n-1)!} F \xi - \frac{\phi^{n-1}}{(n-1)!} \psi \wedge db, \]
\[ W_{n,1,3} = \frac{\phi^{n-3}}{6(n-3)!} \psi \wedge \psi \wedge \psi \xi + \frac{\phi^{n-1}}{(n-1)!} F \wedge db \]
\[ + \frac{\phi^{n-2}}{2(n-2)!} (2\psi \wedge F \xi + \psi \wedge \psi \wedge db). \] (3.106)

These relations correspond to the cocycles \[ \text{Baulieu and Grossman (1988)} \] in \( U(1) \) case.

Next, let us look for the observables for matter sector. The BRST transformation rules in this sector is given by \( \delta_B, \delta_B N, \delta_B C \) and \( \delta_B \phi \). At first sight, the matter sector does not have any observable, but we can find the combined form
\[ \tilde{W} = i\phi \mathcal{M} M + \mathcal{N} N \] (3.107)
is an observable. However, unfortunately, as \( \tilde{W} \) is cohomologically trivial because \( \delta_B \tilde{W} = 0 \) but \( d\tilde{W} \neq \delta_B \tilde{W}' \) for some \( \tilde{W}' \). Accordingly, \( W \) does not give any new topological invariant \[ \text{Ohta (1998)} \].

In topological Bogomol’nyi theory, there is a sequence of observables associated with a magnetic charge. For the Abelian case, it is given by
\[ W = \int_Y F \wedge db. \] (3.108)

As is pointed out for the case of Bogomol’nyi monopoles \[ \text{Birmingham et al. (1989)} \], we cannot get the observables related with this magnetic charge by the action of \( \delta_B \) as well, but we can construct those observables by anti-BRST variation \( \tilde{\delta}_B \) which maps \( (m, n) \)–form to \( (m, n - 1) \)–form. \( \tilde{\delta}_B \) can be obtained by a discrete symmetry which is realized as ‘time reversal symmetry’ in 4D. In our 3D theory, the discrete symmetry is given by
\[ \phi \rightarrow -\lambda, \quad \lambda \rightarrow -\phi, \quad N \rightarrow i\sqrt{2} \mu, \quad \mu \rightarrow \frac{i}{\sqrt{2}} N, \]
\[ \psi_i \rightarrow \chi_i \sqrt{2}, \quad \chi_i \rightarrow \sqrt{2} \psi_i, \quad \eta \rightarrow \sqrt{2} \xi, \quad \xi \rightarrow -\eta \sqrt{2} \] (3.109)
with \[ b \rightarrow -b, \] (3.110)
where \[ (3.110) \] represents an additional symmetry \[ \text{Birmingham et al. (1989)} \]. Note that we must also change \( N \) and \( \mu \) (and their conjugates).
The positive chirality condition for $M$ should be used in order to check the invariance of the action. In this way, we can get anti–BRST transformation rule by substituting (3.109) and (3.110) into (3.98) and then we can get the observables associated with the magnetic charge by using the action of this anti–BRST variation [Birmingham et al. (1989)].

The topological observables available in this theory are the same with those of topological Bogomol’nyi monopoles.

Finally, let us briefly comment on our 3D theory. First note that Lagrangian $L$ and Hamiltonian $H$ in dimensional reduction can be considered as equivalent. This is because the relation between them is defined by $H = p\dot{q} - L$, where $q$ is any field, the overdot means time derivative and $p$ is a canonical conjugate momentum of $q$, and the dimensional reduction requires the time independence of all fields, thus $H = -L$ in this sense. Though we have constructed the three dimensional action directly from the 3D monopole equations, our action may be interpreted essentially as the Hamiltonian of the four dimensional SW theory.

3.2.4.3 Non–Abelian Case

It is easy to extend the results obtained in the previous subsection to non–Abelian case. In this subsection, we summarize the results for the non–Abelian 3D monopoles (for details, see [Ohta (1998)] and references therein).

Non–Abelian Topological Action

With the auxiliary fields $G^a_{\mu\nu}$ and $\nu$, we consider

$$S_c = \frac{1}{2} \int_Y d^3x \left[ (G^a_i - K_i^a)^2 + |\nu - i\gamma^i D_i M - \gamma^0 b M|^2 \right],$$

where

$$K_i^a = \partial b^a + f_{abc} A_i^b b^c - \frac{1}{2} \epsilon_{ijk} F_{ijk}^a + i \bar{M} \sigma_0 T^a M.$$

Note that the minimum of (3.111) with the gauge $G^a_i = \nu = 0$ are given by the non–Abelian 3D monopoles. We take the generator of Lie algebra in the fundamental representation, e.g., for $SU(n)$,

$$(T_a)_{ij}(T^a)^{kl} = \delta_{il} \delta_{jk} - \frac{1}{n} \delta_{ij} \delta_{kl}. $$

Extension to other Lie algebra and representation is straightforward.
The gauge transformation rule for (3.111) is given by

$$
\delta A^a_i = \partial_i \theta^a + f_{abc} A^b_i \theta^c + \epsilon^a_i, \quad \delta b^a = f_{abc} b^b \theta^c + \tau^a, \quad \delta M = i \theta M + \varphi,
$$

$$
\delta G^a_i = f_{abc} G^b_i \theta^c + \left[ -\epsilon_{ijk} (\partial_j \epsilon^{abk} + f_{abc} \epsilon^{ haili} A^h) 
+ \partial_i \tau^a + f_{abc} (\epsilon^{ haili} b^c - \tau^h A^h) \right] + i(\varphi \sigma_0 T^a M + \overline{M} \sigma_0 T^a \varphi),
$$

$$
\delta \nu = i \gamma^i D_i \varphi + \gamma^i \epsilon_i M + \gamma^0 b \varphi + \gamma^0 \tau M + i \theta \nu. \quad (3.112)
$$

Note that we have a $G^a_i$ term in the transformation of $G^a_i$, while it did not appear in Abelian theory.

The gauge algebra (3.112) possesses on–shell zero modes as in the Abelian case. Setting

$$\theta^a = \Lambda^a, \quad \epsilon^a_i = -\partial_i \Lambda^a - f_{abc} A^b_i \Lambda^c, \quad \tau^a = -f_{abc} b^b \Lambda^c, \quad \varphi = -i \Lambda M,$$

we can easily find that (3.112) closes, i.e.,

$$
\delta A^a_i = 0, \quad \delta b^a = 0, \quad \delta M = 0,
$$

$$
\delta G^a_i = f_{abc} G^b_i \Lambda^c [G^b_i - K^b_i]_{on-shell} = 0,
$$

$$
\delta \nu = i \Lambda [\nu - i(\gamma^i D_i - \gamma^0 b)M]_{on-shell} = 0, \quad (3.113)
$$

when the equations of motion of $G^a_i$ and $\nu$ are used. Note that we must use both equations of motion of $G^a_i$ and $\nu$ in the non–Abelian case, while only ‘$\nu$’ was needed for the Abelian theory. Furthermore, as $\varphi$ is a parameter in the spinor space, $\varphi$ is not $g$–valued, in other words, $\varphi \neq \varphi^a T^a$. (3.112) is first stage reducible.

The assortment of ghost fields, the minimal set $\Phi_{min}$ of the fields and the ghost number and the Grassmann parity, furthermore those for $\Phi^*_{min}$ would be obvious.

Then the solution to the master equation will be

$$
S(\Phi_{min}, \Phi^*_{min}) = S_c + \int_Y \text{Tr}(\Delta S_n) d^3 x, \quad \text{where}
$$
\[ \Delta S_n = \]
\[ A^*_n (D^t c + \psi^t) + b^*(i[b, c] + \xi) + M^*(i\epsilon M + N) + \overline{M}^* (-i\epsilon \overline{M} + \overline{N}) \]
\[ + G^*_n \tilde{G}^n - iN^* (\phi M + cN) + i\overline{N}^* (\phi \overline{M} + c\overline{N}) + \nu^* (i\nu + i\gamma^t D_t N + \gamma^t \psi_i M + \gamma^0 b N + \gamma^0 \xi M) \]
\[ + 2i\nu^* \overline{\psi}_i (-D^t \phi - i[\psi_i, c]) + c^* \left( \phi - i\frac{1}{2} \{c, c\} \right) \]
\[ - i\phi^*[\phi, c] - i\frac{1}{2} \{G^*_n, G^{**}\} \phi + i\xi^*[b, \phi] - \{\xi, c\}. \]

Here

\[ \tilde{G}_i = i[c, G_i] - \epsilon_{ijk} D^j \psi^k + D_t \xi + [\psi_i, \xi] + i(\overline{N} \sigma_{i0} T_a T^a M + \overline{M} \sigma_{i0} T_a T^a N). \]

The equations

\[ -\epsilon_{ijk} D^j \psi^k + D_t \xi + [\psi_i, \xi] + i(\overline{N} \sigma_{i0} T_a T^a M + \overline{M} \sigma_{i0} T_a T^a N) = 0, \]
\[ i\gamma^t D_t N + \gamma^t \psi_i M + \gamma^0 b N + \gamma^0 \xi M = 0, \]

can be seen as linearizations of non–Abelian 3D monopoles.

We augment \( \Phi_{\text{min}} \) by new fields \( \chi_a, d_a^i, \mu(\overline{\mu}), \xi(\overline{\xi}), \lambda, \rho, \eta, e \) and the corresponding anti–fields, but Lagrange multipliers fields \( d_a^i, \xi(\overline{\xi}), c, \eta \) are assumed not to have anti–fields for simplicity and therefore their BRST transformation rules are set to zero. This simplification means that we do not take into account of BRST exact terms. In this sense, the result to be obtained will correspond to those of the dimensionally reduced version of the 4D theory up to these terms, i.e., topological numbers.

From the gauge fixing condition

\[ G_i^a = 0, \quad \nu = 0, \quad \partial^t A_i = 0, \quad -D^i \psi_i + i\frac{1}{2} (\overline{N} M - M \overline{N}) = 0, \]

the gauge fermion will be

\[ \Psi = -\chi^t G_i - \overline{\mu} \nu - \mu \overline{\nu} + \rho \partial^t A_i - \lambda \left[ -D^i \psi_i + i\frac{1}{2} (\overline{N} M - M \overline{N}) \right]. \]
The anti–fields are then given by

\begin{align*}
G_i^* &= -\chi_i, \quad \chi_i^* = -G_i, \\
\nu^* &= -\mu, \quad \bar{\nu} = -\nu, \quad \mu = -\bar{\nu}, \\
M^* &= -\frac{i}{2}\lambda N, \quad \bar{M} = \frac{i}{2}\lambda N, \quad N^* = \frac{i}{2}\lambda \bar{M}, \quad \bar{N} = -\frac{i}{2}\lambda M,
\end{align*}

\begin{align*}
\rho^* &= \partial^i A_i, \quad A^*_i = -\partial_i \rho + i[\lambda, \psi_i], \\
\bar{b}^* &= c^* = \xi^* = \phi^* = \zeta^*(\zeta) = 0,
\end{align*}

\begin{align*}
\lambda^* &= -\left[ -D_i \psi^i + [b, \xi] + \frac{i}{2}(NM - \bar{M}N) \right], \quad \psi^*_i = -D_i \lambda.
\end{align*}

Therefore we find the quantum action

\[ S_q = S_c + \int_Y \text{Tr} \left( \tilde{\Delta} S_n \right) d^3 x, \quad \text{where} \quad (3.114) \]

\[ \tilde{\Delta} S_n = -[-D_i \psi^i + [b, \xi] + \frac{i}{2}(NM - \bar{M}N)] \eta - \lambda(D_i D^i \phi) \\
+ iD_i \{ \psi^i, c \} + i\lambda \{ \psi_i, D^i c + \psi^i \} + (\phi \bar{M}M - i\bar{N}N)\lambda \\
- \chi^i [i[c, G_i] + \epsilon_{ijk} D^j \psi^k + D_k \xi + [\psi_k, \xi] + \frac{i}{2}(\bar{N}\sigma^{ij} T_a T^a M + \bar{M}\sigma^{ij} T_a T^a N)] \\
- \bar{\eta}(i\gamma^\mu D_\mu N + \gamma^\mu \psi_\mu \bar{M} + ic\nu) + (i\gamma^i D_i \bar{N} + \gamma^\mu \psi_\mu \bar{M} + ic\nu)\mu \\
+ 2i\omega \bar{\eta} \mu - \frac{i}{2} \{ \chi_i, \chi^i \} \phi + \rho(\partial_i D^i c + \partial_i \psi^i) - d^i G_i - \bar{\zeta} \nu - \bar{\nu} \zeta + e\partial_i A_i.
\]

In this quantum action, setting

\[ M(\bar{M}) = N(\bar{N}) = \mu(\bar{\eta}) = \nu(\bar{\nu}) = 0, \]

we can find that the resulting action coincides with that of Bogomol’nyi monopoles \cite{Birmingham et al. (1989)}.

Finally, in order to get the off–shell quantum action, both the auxiliary fields should be integrated out by the similar technique presented in Abelian case.
BRST transformation

The BRST transformation rule is given by

\[ \delta_B A_i = -\epsilon (D_i c + \psi_i), \quad \delta_B b = -\epsilon (i[c, b] + \xi), \quad \delta_B \xi = i\epsilon (\{b, \phi\} - \{\xi, c\}), \]
\[ \delta_B M = -\epsilon (i\epsilon M + N), \quad \delta_B G_i = -\epsilon (\tilde{G}_i - i[X_i, \phi]), \]
\[ \delta_B \psi_i = -\epsilon (i\nu + \gamma^\rho D_\rho N + \gamma^\rho \psi_\rho M - i\mu \phi), \quad \delta_B c = \epsilon (\phi - \frac{i}{2} \{c, c\}), \]
\[ \delta_B M = -\epsilon (i\epsilon M + N), \quad \delta_B G_i = -\epsilon (\tilde{G}_i - i[X_i, \phi]), \]
\[ \delta_B \psi_i = -\epsilon (D_i \phi + i\{\psi_i, c\}), \quad \delta_B \rho = \epsilon c, \quad \delta_B \lambda = -c\eta, \]
\[ \delta_B M = -\epsilon (i\nu + \gamma^\rho D_\rho N + \gamma^\rho \psi_\rho M - i\mu \phi), \quad \delta_B c = \epsilon (\phi - \frac{i}{2} \{c, c\}), \]
\[ \delta_B M = -\epsilon (i\nu + \gamma^\rho D_\rho N + \gamma^\rho \psi_\rho M - i\mu \phi), \quad \delta_B \rho = \epsilon c, \quad \delta_B d_i = \delta_B e = \delta_B \zeta = \delta_B \eta = 0. \quad (3.115) \]

It is easy to get supersymmetry also in this case. However, as we have omitted the BRST exact terms, the supersymmetry in our construction does not detect them.

Observables

We have already constructed the topological observables for Abelian case. Also in non-Abelian case, the construction of observables is basically the same. But the relation (3.101) and (3.102) are required to modify

\[ (d + \delta_B)A - \frac{i}{2} [A, A] = \mathcal{F}, \quad (d + \delta_B)b - i\epsilon [A, b] = \mathcal{K}, \quad (3.116) \]
\[ (d + \delta_B)\mathcal{F} - i[A, \mathcal{F}] = 0, \quad (d + \delta_B)\mathcal{K} - i[A, \mathcal{K}] = i[\mathcal{F}, b], \quad (3.117) \]

respectively, where \([*, *]\) is a graded commutator. The observables in geometric and matter sector are the same as before, but we should replace \(db\) by \(dA\) in (3.106) as well as (3.116) and (3.117), where \(dA\) is an exterior covariant derivative and trace is required. In addition, the magnetic charge observables are again obtained by anti-BRST variation as outlined before.

The observables in geometric sector are those in (3.104) and follow the cohomological relation (3.105). In this way, the topological observables available in this three dimensional theory are precisely the Bogomol’nyi monopole cocycles [Baulieu and Grossman (1988)].
3.3 Stringy Geometrodynamics

3.3.1 Stringy Actions and Amplitudes

Now we give a brief review of modern path-integral methods in superstring theory (mainly following [Deligne et al. (1999)]). Recall that the fundamental quantities in quantum field theory (QFT) are the transition amplitudes $\text{Amp} : \text{IN} \rightarrow \text{OUT}$, describing processes in which a number $\text{IN}$ of incoming particles scatter to produce a number $\text{OUT}$ of outgoing particles. The square modulus of the transition amplitude yields the probability for this process to take place.

3.3.1.1 Strings

Recall that in string theory, elementary particles are not described as 0-dimensional points, but instead as 1D strings. If $\mathcal{M}_s$ and $\mathcal{M} (\sim \mathbb{R} \times \mathcal{M}_s)$ denote the 3D space and 4D space–time manifolds respectively, then we picture strings as in Figure 3.11.

![Figure 3.11](image)

Fig. 3.11 Basic geometrical objects of string theory: (a) a space with fixed time; (b) a space–time picture; (c) a point–particle; (d) a world–line of a point–particle; (e) a closed string; (f) a world–sheet of a closed string; (g) an open string; (h) a world–sheet of an open string.

While the point–particle sweeps out a 1D world–line, the string sweeps out a world–sheet, i.e., a 2D real surface. For a free string, the topology of the world–sheet is a cylinder (in the case of a closed string) or a sheet (for an open string).

Roughly, different elementary particles correspond to different vibration
modes of the string just as different minimal notes correspond to different vibrational modes of musical string instruments.

It turns out that the physical size of strings is set by gravity, more precisely the Planck length $\ell_P \sim 10^{-33}$ cm. This scale is so small that we effectively only see point–particles at our distance scales. Thus, for length scales much larger than $\ell_P$, we expect to recover a QFT–description of point–particles, plus typical string corrections that represent physics at the Planck scale.

3.3.1.2 Interactions

While the string itself is an extended 1D object, the fundamental string interactions are local, just as for point–particles. The interaction takes place when strings overlap in space at the same time. In case of closed string theories the interactions have a form depicted in Figure 3.12, while in case of open string theories the interactions have a form depicted in Figure 3.13. Other interactions result from combining the interactions defined above.

Fig. 3.12 Interactions in closed string theories (left 2D–picture and right 3D–picture).

Fig. 3.13 Interactions in open string theories (left 2D–picture and right 3D–picture).

In point–particle theories, the fundamental interactions are read off from the QFT–Lagrangian. An interaction occurs at a geometrical point, where
the world–lines join and cease to be a manifold. In Lorentz–invariant theories (where manifold $M$ is a flat Minkowski space–time), the interaction point is Lorentz–invariant. To specify how the point–particles interact, additional data must be supplied at the interaction point, giving rise to many possible distinct quantum field theories.

In string theory, the interaction point depends upon the Lorentz frame chosen to observe the process. In the Figure above, equal time slices are indicated from the point of view of two different Lorentz frames, schematically indicated by $t$ and $t'$. The closed string interaction, as seen from frames $t$ and $t'$, occurs at times $t_2$ and $t'_2$ and at (distinct) points $P$ and $P'$ respectively.

Lorentz invariance of interaction forbids that any point on the world–sheet be singled out as interaction point. Instead, the interaction results purely from the joining and splitting of strings. While free closed strings are characterized by their topology being that of a cylinder, interacting strings are characterized by the fact that their associated world–sheet is connected to at least 3 strings, incoming and/or outgoing.

As a result, the free string determines the nature of the interactions completely, leaving only the string coupling constant undetermined.

The orientation is an additional structure of closed strings, dividing them into two categories: (i) oriented strings, in which all world–sheets are assumed to be orientable; and (ii) non–oriented strings, in which world–sheets are non–orientable, such as the Möbius strip, Klein bottle, etc.

### 3.3.1.3 Loop Topology of Closed Surfaces

For simplicity, here we consider closed oriented strings only, so that the associated world–sheet is also oriented. A general string configuration describing the process in which $M$ incoming strings interact and produce $N$ outgoing strings looks at the topological level like a closed surface with $M + N = E$ boundary components and any number of handles (see Figure 3.14). This picture is a kind of topological generalization of nonlinear control MIMO–systems with $M$ inputs, $N$ outputs $X$ states.

The internal loops may arise when virtual particle pairs are produced, just as in quantum field theory. For example, a Feynman diagram in quantum field theory that involves a loop is shown in Figure 3.15 together with the corresponding string diagram.

Surfaces associated with closed oriented strings have two topological invariants: (i) the number of boundary components $E = M + N$ (which may
Fig. 3.14 Boundary components and handles of closed oriented system of $M$ incoming strings, interacting through internal loops, to produce $N$ outgoing strings. Note the striking similarity with MIMO–systems of nonlinear control theory, with $M$ input processes and $N$ output processes.

Fig. 3.15 A QFT Feynman diagram that involves an internal loop (left), with the corresponding string diagram (right).

be shrunk to punctures, under certain conditions), and (ii) the number $h$ of handles on the surface, which equals the surface genus.

Fig. 3.16 Number $h$ of handles on the surface of closed oriented strings, which equals the string–surface genus: (a) $h = 0$ for sphere $S^2$; (b) $h = 1$ for torus $T^2$; (c) $h = 2$ for string–surfaces with higher genus, etc.

When $E = 0$, we just have the topological classification of compact ori-
ented surfaces without boundary. Rendering $E > 0$ is achieved by removing $E$ discs from the surface.

Recall that in QFT, an expansion in powers of Planck’s constant $\hbar$ yields an expansion in the number of loops of the associated Feynman diagram, for a given number of external states:

$$\hbar^{E+h-1} = \begin{cases} \frac{\hbar}{h} & \text{for every propagator} \\ \frac{1}{h} \hbar^{-1} & \text{for every vertex} \\ -1 & \text{for overall momentum conservation} \end{cases}$$

Thus, in string theory we expect that, for a given number of external strings $E$, the topological expansion genus by genus should correspond to a loop expansion as well.

Recall that in QFT, there are in general many Feynman diagrams that correspond to an amplitude with a given number of external particles and a given number of loops. For example, for $E = 4$ external particles and $h = 1$ loop in $\phi^3$ theory are given in Figure 3.17, together with the same process in string theory (for closed oriented strings), where it is described by just a single diagram (right).

![Fig. 3.17 Feynman QFT–diagrams for $\phi^3$ theory with $E = 4$ external particles and $h = 1$ loop (left), and a single corresponding string diagram (right). In this way the usual Feynman diagrams of quantum field theory are generalized by arbitrary Riemannian surfaces.](image)

Much of recent interest has been focused on the so-called $D$–branes. A $D$–brane is a submanifold of space–time with the property that strings can end or begin on it.

### 3.3.2 Transition Amplitudes for Strings

The only way we have today to define string theory is by giving a rule for the evaluation of transition amplitudes, order by order in the loop expansion, i.e., genus by genus. The rule is to assign a relative weight to a given
configuration and then to sum over all configurations [Deligne et al. (1999)]. To make this more precise, we first describe the system’s configuration manifold $M$ (see Figure 3.18).

Fig. 3.18 The embedding map $x$ from the reference surface $\Sigma$ into the pseudo–Riemannian configuration manifold $M$ (see text for explanation).

We assume that $\Sigma$ and $M$ are smooth manifolds, of dimensions 2 and $n$ respectively, and that $x$ is a continuous map from $\Sigma$ to $M$. If $\xi^m$, (for $m = 1, 2$), are local coordinates on $\Sigma$ and $x^\mu$, ($\mu = 1, \ldots, n$), are local coordinates on $M$ then the map $x$ may be described by functions $x^\mu(\xi^m)$ which are continuous.

To each system configuration we can associate a weight $e^{-S[x, \Sigma, M]}$, (for $S \in \mathbb{C}$) and the transition amplitude $\text{Amp}$ for specified external strings (incoming and outgoing) is get by summing over all surfaces $\Sigma$ and all possible maps $x$,

$$Amp = \sum_{\text{surfaces } \Sigma} \sum_{x} e^{-S[x, \Sigma, M]}.$$

We now need to specify each of these ingredients:

1. We assume $M$ to be an $n$D Riemannian manifold, with metric $g$. A special case is flat Euclidean space–time $\mathbb{R}^n$. The space–time metric is assumed fixed.

$$ds^2 = (dx, dx)_g = g_{\mu\nu}(x) dx^\mu dx^\nu.$$

2. The metric $g$ on $M$ induces a metric on $\Sigma$: $\gamma = x^*(g)$,

$$\gamma = \gamma_{mn} d\xi^m \otimes d\xi^n, \quad \gamma_{mn} = g_{\mu\nu} \frac{\partial x^\mu}{\partial \xi^m} \frac{\partial x^\nu}{\partial \xi^n}.$$
This metric is non-negative, but depends upon $x$. It is advantageous to introduce an intrinsic Riemannian metric $g$ on $\Sigma$, independently of $x$; in local coordinates, we have

$$g = g_{mn}(\xi) \, d\xi^m \, d\xi^n.$$  

A natural intrinsic candidate for $S$ is the area of $x(\Sigma)$, which gives the so-called Nambu–Goto action\(^{17}\)

$$\text{Area} (x(\Sigma)) = \int_\Sigma d\mu_\gamma = \int_\Sigma n^2 \xi \sqrt{\det \gamma_{mn}},$$  \tag{3.118}

which depends only upon $g$ and $x$, but not on $g$ \cite{Goto1971}. However, the transition amplitudes derived from the Nambu–Goto action are not well–defined quantum–mechanically.

Otherwise, we can take as starting point the so–called Polyakov action\(^{18}\)

$$S[x,g] = \kappa \int_\Sigma (dx, *dx)_g = \kappa \int_\Sigma d\mu_y g^{mn} \partial_m x^\mu \partial_n x^\nu \gamma_{\mu\nu}(x),$$  \tag{3.119}

where $\kappa$ is the string tension (a positive constant with dimension of inverse length square). The stationary points of $S$ with respect to $g$ are at $g^0 = e^{\phi} \gamma$ for some function $\phi$ on $\Sigma$, and thus $S[x,g^0] \sim \text{Area} (x(\Sigma))$.

The Polyakov action leads to well–defined transition amplitudes, get by integration over the space $\text{Met}(\Sigma)$ of all positive metrics on $\Sigma$ for a given topology, as well as over the space of all maps $\text{Map}(\Sigma,M)$. We can define the path integral

$$\text{Amp} = \sum_{\text{topologies}} \int_{\text{Met}(\Sigma)} \frac{1}{N(g)} \int_{\text{Map}(\Sigma,M)} \mathcal{D}[x] \, e^{-S[x,g,g]},$$

where $N$ is a normalization factor, while the measures $\mathcal{D}[g]$ and $\mathcal{D}[x]$ are constructed from $\text{Diff}^+(\Sigma)$ and $\text{Diff}(M)$ invariant $L^2$ norms on $\Sigma$ and $M$. For fixed metric $g$, the action $S$ is well–known: its stationary points are the harmonic maps $x: \Sigma \to M$ (see, e.g., \cite{EellsLemaire1978}). However, $g$ here varies and in fact is to be integrated over. For a general metric $g$, the action $S$ defines a nonlinear sigma model, which is renormalizable.

\(^{17}\)Nambu–Goto action is the starting point of the analysis of string behavior, using the principles of ordinary Lagrangian mechanics. Just as the Lagrangian for a free point particle is proportional to its proper time—i.e., the 'length' of its world–line, a relativistic string’s Lagrangian is proportional to the area of the sheet which the string traces as it travels through space–time.

\(^{18}\)The Polyakov action is the 2D action from conformal field theory, used in string theory to describe the world–sheet of a moving string.
because the dimension of $\Sigma$ is 2. It would not in general be renormalizable in dimension higher than 2, which is usually regarded as an argument against the existence of fundamental membrane theories (see Deligne et al. (1999)).

The Nambu–Goto action (3.118) and Polyakov action (3.119) represent the core of the so-called bosonic string theory, the original version of string theory, developed in the late 1960s. Although it has many attractive features, it also predicts a particle called the tachyon possessing some unsettling properties, and it has no fermions. All of its particles are bosons, the matter particles. The physicists have also calculated that bosonic string theory requires 26 space-time dimensions: 25 spatial dimensions and one dimension of time. In the early 1970s, supersymmetry was discovered in the context of string theory, and a new version of string theory called superstring theory (i.e., supersymmetric string theory) became the real focus, as it includes also fermions, the force particles. Nevertheless, bosonic string theory remains a very useful ‘toy model’ to understand many general features of perturbative string theory.

3.3.3 Weyl Invariance and Vertex Operator Formulation

The action $S$ is also invariant under Weyl rescalings of the metric $g$ by a positive function on $\sigma: \Sigma \rightarrow \mathbb{R}$, given by $g \rightarrow e^{2\sigma} g$. In general, Weyl invariance of the full amplitude may be spoiled by anomalies. Assuming Weyl invariance of the full amplitude, the integral defining $Amp$ may be simplified in two ways.

1) The integration over $Met(\Sigma)$ effectively collapses to an integration over the moduli space of surfaces, which is finite dimensional, for each genus $h$. 

2) The boundary components of $\Sigma$ — characterizing external string states — may be mapped to regular points on an underlying compact surface without boundary by conformal transformations. The data, such as momenta and other quantum numbers of the external states, are mapped into vertex operators. The amplitudes are now given by the path integral

$$Amp = \sum_{h=0}^{\infty} \int_{\text{Met}(\Sigma)} \mathcal{D}[g] \frac{1}{N(g)} \int_{\text{Map}(\Sigma,M)} \mathcal{D}[x] V_1 \ldots V_N e^{-S},$$

for suitable vertex operators $V_1, \ldots, V_N$. 

3.3.4 More General Stringy Actions

Generalizations of the action $S$ given above are possible when $M$ carries extra structure.

1) $M$ carries a 2–form $B \in \Omega(2)(M)$. The resulting contribution to the action is also that of a ‘nonlinear sigma model’

$$S_B[x, B] = \int_{\Sigma} dx^\nu B_{\mu\nu}(x).$$

2) $M$ may carry a dilaton field $\Phi \in \Omega(0)(M)$ so that

$$S_\Phi[x, \Phi] = \int_{\Sigma} d\mu g^R g\Phi(x).$$

where $R_g$ is the Gaussian curvature of $\Sigma$ for the metric $g$.

3) There may be a tachyon field $T \in \Omega(0)(M)$ contributing

$$S_T[x, T] = \int_{\Sigma} d\mu gT(x).$$

3.3.5 Transition Amplitude for a Single Point Particle

The transition amplitude for a single point–particle could in fact be get in a way analogous to how we prescribed string amplitudes. Let space–time be again a Riemannian manifold $M$, with metric $g$. The prescription for the transition amplitude of a particle travelling from a point $y \in M$ to a point $y'$ to $M$ is expressible in terms of a sum over all (continuous) paths connecting $y$ to $y'$:

$$\text{Amp}(y, y') = \sum_{\text{paths joining } y \text{ and } y'} e^{-S[\text{path}]}.$$

Paths may be parametrized by maps from $C = [0, 1]$ into $M$ with $x(0) = y$, $x(1) = y'$. A simple world–line action for a massless particle is get by introducing a metric $g$ on $[0, 1]$

$$S[x, g] = \frac{1}{2} \int_C d\tau g(\tau)^{-1} \dot{x}^\mu \dot{x}^\nu g_{\mu\nu}(x),$$

which is invariant under $Diff^+(C)$ and $Diff(M)$. 
Recall that the analogous prescription for the point–particle transition amplitude is the path integral

\[ \text{Amp}(y, y') = \int_{\text{Met}(C)} \mathcal{D}[g] \frac{1}{N} \int_{\text{Map}(C, M)} \mathcal{D}[x] e^{-S[x, g]}. \]

Note that for string theory, we had a prescription for transition amplitudes valid for all topologies of the world–sheet. For point–particles, there is only the topology of the interval \( C \), and we can only describe a single point–particle, but not interactions with other point–particles. To put those in, we would have to supply additional information.

Finally, it is very instructive to work out the amplitude \( \text{Amp} \) by carrying out the integrations. The only \( \text{Diff}^+(C) \) invariant of \( g \) is the length \( L = \int_0^1 d\tau g(\tau) \); all else is generated by \( \text{Diff}^+(C) \). Defining the normalization factor to be the volume of \( \text{Diff}(C) \): \( N = \text{Vol}(\text{Diff}(C)) \) we have \( \mathcal{D}[g] = \mathcal{D}[v] dL \) and the transition amplitude becomes

\[ \text{Amp}(y, y') = \int_0^\infty dL \int \mathcal{D}[x] e^{-\frac{1}{2} \int_0^1 d\tau (\dot{x}, \dot{x})_0} = \int_0^\infty dL \langle y' | e^{-L\Delta} | y \rangle = \langle y' | \frac{1}{\Delta} | y \rangle. \]

Thus, the amplitude is just the Green function at \( (y, y') \) for the Laplacian \( \Delta \) and corresponds to the propagation of a massless particle (see \cite{Deligne et al. (1999)}).

### 3.3.6 Witten’s Open String Field Theory

Noncommutative nature of space–time has often appeared in non–perturbative aspects of string theory. It has been used in a formulation of interacting open string field theory by Ed Witten \cite{Witten (1986), Witten (1988c)}. Witten has written a classical action of open string field theory in terms of noncommutative geometry, where the noncommutativity appears in a product of string fields. Later, the \textit{Dirichlet branes} (or, D–branes) have been recognized as solitonic objects in superstring theory \cite{Polchinski (1995)}. Further, it has been found that the low energy behavior of the D–branes are well described by supersymmetric Yang–Mills theory (SYM) \cite{Witten (1996)}. In the situation of some D–branes coinciding, the space–time coordinates are promoted to matrices which appear as the fields in SYM. Then the size of the matrices corresponds to the num-
number of the D–branes, so noncommutativity of the matrices is related to the noncommutative nature of space–time.

In this subsection, mainly following [Sugino (2000)], we review some basic properties of Witten’s bosonic open string field theory [Witten (1988c)] and its explicit construction based on a Fock space representation of string field functional and $\delta$–function overlap vertices [Gross and Jevicki (1987a); Gross and Jevicki (1987b); Cremmer et al. (1986)].

Witten introduced a beautiful formulation of open string field theory in terms of a noncommutative extension of differential geometry, where string fields, the BRST operator $Q$ and the integration over the string configurations $\int$ in string field theory are analogs of differential forms, the exterior derivative $d$ and the integration over the manifold $\int_M$ in the differential geometry, respectively. The ghost number assigned to the string field corresponds to the degree of the differential form. Also the (noncommutative) product between the string fields $*$ is interpreted as an analog of the wedge product $\wedge$ between the differential forms.

The axioms obeyed by the system of $\int$, $*$ and $Q$ are

$$\int Q A = 0, \quad Q(A*B) = (QA)*B + (-1)^{n_A} A*(QB),$$
$$A*B*C = A*(B*C), \quad \int A*B = (-1)^{n_A n_B} \int B*A,$$

where $A$, $B$ and $C$ are arbitrary string fields, whose ghost number is half–integer valued: The ghost number of $A$ is defined by the integer $n_A$ as $n_A + \frac{1}{2}$.

Then Witten discussed the following string–field–theory action

$$S = \frac{1}{G_s} \int \left( \frac{1}{2} \psi*Q\psi + \frac{1}{3} \psi^* \psi * \psi \right),$$

(3.120)

where $G_s$ is the open string coupling constant and $\psi$ is the string field with the ghost number $-\frac{1}{2}$. The action is invariant under the gauge transformation

$$\delta\psi = QA + \psi*A - A*\psi,$$

with the gauge parameter $A$ of the ghost number $-\frac{3}{2}$.

3.3.6.1 Operator Formulation of String Field Theory

The objects defined above can be explicitly constructed by using the operator formulation, where the string field is represented as a Fock space,
and the integration $\int$ as an inner product on the Fock space. It was considered by [Gross and Jevicki (1987a) Gross and Jevicki (1987b)] in the case of the Neumann boundary condition. We will heavily use the notation of [Gross and Jevicki (1987a) Gross and Jevicki (1987b)]. In the operator formulation, the action (3.120) is described as

$$S = \frac{1}{G_s} \left( \frac{1}{2} |\langle V_2 | \psi \rangle |^2 Q |\langle \psi \rangle \rangle_2 + \frac{1}{3} |\langle V_3 | \psi \rangle |^2 |\langle \psi \rangle \rangle_2 |\langle \psi \rangle \rangle_3 \right), \quad (3.121)$$

where the structure of the product $*$ in the kinetic and potential terms is encoded to that of the overlap vertices $\langle V_2 |$ and $\langle V_3 |$ respectively (here, subscripts put to vectors in the Fock space label the strings concerning the vertices).

As a preparation for giving the explicit form of the overlaps, let us consider open strings in 26-dimensional space–time with the constant metric $G_{ij}$ in the Neumann boundary condition. The world sheet action is given by

$$S_{WS} = \frac{1}{4\pi\alpha'} \int d\tau \int_0^\pi d\sigma G_{ij} (\partial_\tau X^i \partial_\tau X^j - \partial_\sigma X^i \partial_\sigma X^j) + S_{gh}, \quad (3.122)$$

where $S_{gh}$ is the action of the $bc$–ghosts:

$$S_{gh} = \frac{i}{2\pi} \int d\tau \int_0^\pi d\sigma [c_+ (\partial_\tau - \partial_\sigma) b_+ + c_- (\partial_\tau + \partial_\sigma) b_-]. \quad (3.123)$$

Under the Neumann boundary condition, the string coordinates have the standard mode expansions:

$$X^j(\tau, \sigma) = x^j + 2\alpha' \tau p^j + i\sqrt{2\alpha'} \sum_{n \neq 0} \frac{1}{n} \alpha_n^j e^{-in\tau} \cos(n\sigma), \quad (3.124)$$

also the mode expansions of the ghosts are given by

$$c_\pm(\tau, \sigma) = \sum_{n \in \mathbb{Z}} c_n e^{-in(\tau \pm \sigma)} \equiv e(\tau, \sigma) \pm i\pi_0(\tau, \sigma),$$

$$b_\pm(\tau, \sigma) = \sum_{n \in \mathbb{Z}} b_n e^{-in(\tau \pm \sigma)} \equiv \pi_c(\tau, \sigma) \mp ib(\tau, \sigma).$$

As a result of the quantization, the modes obey the commutation relations:

$$[x^i, p^j] = iG^{ij}, \quad [\alpha_n^i, \alpha_m^j] = nG^{ij} \delta_{n+m,0}, \quad \{b_n, c_m\} = \delta_{n+m,0},$$

while the other terms vanish.
The overlap \(|V_N| = |V_N|^X|V_N|^{gh}\), \((N = 1, 2, \cdots)\) is the state satisfying the continuity conditions for the string coordinates and the ghosts at the \(N\)-string vertex of the string field theory. The superscripts \(X\) and \(gh\) show the contribution of the sectors of the coordinates and the ghosts respectively. The continuity conditions for the coordinates are
\[
(X^{(r)}(\sigma) - X^{(r-1)}(\pi - \sigma))|V_N|^X = 0, \quad (P_i^{(r)}(\sigma) + P_i^{(r-1)}(\pi - \sigma))|V_N|^X = 0,
\]
for \(0 \leq \sigma \leq \frac{\pi}{2}\) and \(r = 1, \cdots, N\). Here \(P_i(\sigma)\) is the momentum conjugate to the coordinate \(X^i(\sigma)\) at \(\tau = 0\), and the superscript \((r)\) labels the string \((r)\) meeting at the vertex. In the above formulas, we regard \(r = 0\) as \(r = N\) because of the cyclic property of the vertex. For the ghost sector, we impose the following conditions on the variables \(c(\sigma), b(\sigma)\) and their conjugate momenta \(\pi_c(\sigma), \pi_b(\sigma)\):
\[
(\pi_c^{(r)}(\sigma) - \pi_c^{(r-1)}(\pi - \sigma))|V_N|^{gh} = 0, \quad (b^{(r)}(\sigma) - b^{(r-1)}(\pi - \sigma))|V_N|^{gh} = 0,
\]
\[
(c^{(r)}(\sigma) + c^{(r-1)}(\pi - \sigma))|V_N|^{gh} = 0, \quad (\pi_b^{(r)}(\sigma) + \pi_b^{(r-1)}(\pi - \sigma))|V_N|^{gh} = 0,
\]
for \(0 \leq \sigma \leq \frac{\pi}{2}\) and \(r = 1, \cdots, N\).
on $\sigma = 0, \pi$. This can be rewritten to the convenient form

$$E_{ij} \partial_\pm X^j = (E^T)_{ij} \partial_\pm X^j,$$  \hspace{1cm} (3.128)

where $E_{ij} \equiv g_{ij} + 2\pi\alpha' B_{ij}$, and $\partial_\pm$ are derivatives with respect to the light cone variables $\sigma^\pm = \tau \pm \sigma$.

We can easily see that $X^j(\tau, \sigma)$ satisfying the boundary condition (3.128) has the following mode expansion:

$$X^j(\tau, \sigma = \tilde{x}^j + \alpha' \left[ (E^{-1})^j k_{il} p^l \sigma^- + (E^{-1T})^j k_{il} p^l \sigma^+ \right] \right),$$  \hspace{1cm} (3.129)

We will get the commutators between the modes from the propagator of the open strings, which gives another derivation different from the method by [Chu and Ho (1999)] based on the quantization via the Dirac bracket. When performing the Wick rotation: $\tau \rightarrow -i\tau$ and mapping the world sheet to the upper half plane $z = e^{\tau + i\sigma}, \bar{z} = e^{\tau - i\sigma} (0 \leq \sigma \leq \pi)$, the boundary condition (3.128) becomes

$$E_{ij} \partial_\pm X^j = (E^T)_{ij} \partial_\pm X^j,$$  \hspace{1cm} (3.130)

which is imposed on the real axis $z = \bar{z}$. The propagator $\langle X^i(z, \bar{z})X^j(z', \bar{z}') \rangle$ satisfying the boundary condition (3.130) is determined as

$$\langle X^i(z, \bar{z})X^j(z', \bar{z}') \rangle = -\alpha' \left[ g^{ij} \ln |z - z'| - g^{ij} \ln |z - \bar{z}'| \right] + G^{ij} \ln |z - \bar{z}'|^2 + \frac{1}{2\pi\alpha'} \theta^{ij} \ln \frac{z - \bar{z}'}{z - z'} + D^{ij},$$

where $G^{ij}$ and $\theta^{ij}$ are given by

$$G^{ij} = \frac{1}{2}(E^{-1} + E^{-1})^{ij} = (E^{-1} g E^{-1})^{ij} = (E^{-1} g E^{-1})^{ij}, \hspace{1cm} (3.131)$$

$$\theta^{ij} = 2\pi\alpha' \cdot \frac{1}{2}(E^{-1} - E^{-1})^{ij} = (2\pi\alpha')^2(E^{-1} g E^{-1})^{ij} \hspace{1cm} \quad (3.132)$$

Also the constant $D^{ij}$ remains unknown from the boundary condition alone. However it is an irrelevant parameter, so we can fix an appropriate value.
The mode-expanded form (3.129) is mapped to
\[ X_j(z, \bar{z}) = \tilde{x}_j - i\alpha'[(E - 1)_{jk} p_k \ln \bar{z} + (E^{-1T})_{jk} p_k \ln z] \]
\[ + i\sqrt{2\alpha'} \sum_{n \neq 0} \frac{1}{n} \left[(E - 1)_{jk} \alpha_{n,k} \bar{z}^{-n} + (E^{-1T})_{jk} \alpha_{n,k} \bar{z}^{-n}\right]. \]

Note that the indices of \( p^i \) and \( \alpha_n^j \) were lowered by the metric \( g_{ij} \) not \( G_{ij} \).

Recall the definition of the propagator
\[ \langle X_i(z, \bar{z}) X_j(z', \bar{z}') \rangle \equiv R(X_i(z, \bar{z}) X_j(z', \bar{z}')) - N(X_i(z, \bar{z}) X_j(z', \bar{z}')), \tag{3.133} \]
where \( R \) and \( N \) stand for the radial ordering and the normal ordering respectively. We take a prescription for the normal ordering which pushes \( p_i \) to the right and \( \tilde{x}_j \) to the left with respect to the zero–modes \( p_i \) and \( \tilde{x}_j \).

It corresponds to considering the vacuum satisfying
\[ p_j|0\rangle = \alpha_n^j|0\rangle = 0 \quad (n > 0), \quad |0\rangle_{\alpha_n^j} = 0 \quad (n < 0), \tag{3.134} \]
which is the standard prescription for calculating the propagator of the massless scalar field in 2D conformal field theory from the operator formalism. Making use of (3.133), (3.134) and techniques of the contour integration, it is easy to get the commutators
\[ [\alpha_{n,i}, \alpha_{m,j}] = n\delta_{n+m,0} G_{ij}, \quad [\tilde{x}^i, p_j] = i\delta^i_j, \]
where the first equation holds for all integers with \( \alpha_{0,i} \equiv \sqrt{2\alpha'} p_i \). The constant \( D_{ij} \) is written as \( \alpha' D_{ij} = -\langle 0|\tilde{x}^i \tilde{x}^j|0\rangle \). Let us fix \( D_{ij} \) as \( \alpha' D_{ij} = -\frac{1}{2} \theta_{ij} \), which is the convention taken in [Seiberg and Witten (1999)]. Then the coordinates \( \tilde{x}^i \) become noncommutative:
\[ [\tilde{x}^i, \tilde{x}^j] = i\theta_{ij}, \]
but the center of mass coordinates \( x^i \equiv \tilde{x}^i + \frac{1}{2} \theta^i_j p_j \) can be seen to commute each other.

Now we have the mode–expanded form of the string coordinates and the commutation relations between the modes, which are
\[ X^i(\tau, \sigma) = x^i + 2\alpha' \left( G^{jk} \tau + \frac{1}{2\pi \alpha'} \theta^{jk} (\sigma - \frac{\pi}{2}) \right) p_k \]
\[ + i\sqrt{2\alpha'} \sum_{n \neq 0} \frac{1}{n} e^{-in\tau} \left[G^{jk} \cos(n\sigma) - i\frac{1}{2\pi \alpha'} \theta^{jk} \sin(n\sigma)\right] \alpha_{n,k}, \]
\[ [\alpha_{n,i}, \alpha_{m,j}] = n\delta_{n+m,0} G_{ij}, \quad [x^i, p_j] = i\delta^i_j, \]
with all the other commutators vanishing.

Also, due to the formula
\[ \sum_{n=1}^{\infty} \frac{2}{n} \sin(n(\sigma + \sigma')) = \begin{cases} \pi - \sigma - \sigma', & (\sigma + \sigma' \neq 0, 2\pi) \\ 0, & (\sigma + \sigma' = 0, 2\pi) \end{cases} \]
we can see by a direct calculation that the end points of the string become noncommutative
\[ [X^i(\tau, \sigma), X^j(\tau, \sigma')] = \begin{cases} i\theta_{ij}, & (\sigma = \sigma' = 0) \\ -i\theta_{ij}, & (\sigma = \sigma' = \pi) \\ 0, & (\text{otherwise}) \end{cases} \]

On the other hand, it is noted that the conjugate momenta have the mode expansion identical with that in the Neumann case:
\[ P_i(\tau, \sigma) = \frac{1}{2\pi\alpha'} (g_{ij} \partial_\tau - 2\pi\alpha' B_{ij} \partial_\sigma) X^j(\tau, \sigma) = \frac{1}{\pi} p_i + \frac{1}{\pi\sqrt{2\alpha'}} \sum_{n \neq 0} e^{-in\tau} \cos(n\sigma) \alpha_{n,i}. \]

Note that the relations (3.131) and (3.132) are in the same form as a T–duality transformation, although the correspondence is a formal sense, because we are not considering any compactification of space–time. The generalized T–duality transformation, namely $O(D,D)$–transformation, is defined by
\[ E' = (aE + b)(cE + d)^{-1}, \]
with $a$, $b$, $c$ and $d$ being $D \times D$ real matrices. ($D$ is the dimension of space–time.) The matrix $h = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $O(D,D)$ matrix, which satisfies
\[ h^T J h = J, \quad \text{where} \quad J = \begin{pmatrix} 0 & 1_D \\ 1_D & 0 \end{pmatrix}. \]
The relations (3.131) and (3.132) correspond to the case of the inversion $a = d = 0, b = c = 1_D$.

3.3.6.3 Construction of Overlap Vertices

Here we construct Witten’s open string theory in the constant $B$–field background by obtaining the explicit formulas of the overlap vertices. As is understood from the fact that the action of the ghosts (3.123) contains no
background fields, the ghost sector is not affected by turning on the $B$–field
background. Thus we may consider the coordinate sector only. First, let us
see the mode-expanded forms of the coordinates and the momenta at $\tau = 0$

$$X^i(\sigma) = G^{jk} y_k + \frac{1}{\pi} \theta^{jk} (\sigma - \frac{\pi}{2}) p_k$$

$$+ 2 \sqrt{\alpha'} \sum_{n=1}^{\infty} \left[ G^{jk} \cos(n\sigma)x_{n,k} + \frac{1}{2\pi \alpha'} \theta^{jk} \sin(n\sigma) \frac{1}{n} p_{n,k} \right],$$

$$P_i(\sigma) = \frac{1}{\pi} p_i + \frac{1}{\pi \sqrt{\alpha'}} \sum_{n=1}^{\infty} \cos(n\sigma)p_{n,i},$$

where $x^j = G^{jk} y_k$, the coordinates and the momenta for the oscillator
modes are

$$x_{n,k} = \frac{i}{2} \sqrt{\frac{2}{n}} (a_{n,k} - a_{n,k}^+) = \frac{i}{\sqrt{2n}} (\alpha_{n,k} - \alpha_{-n,k}),$$

$$p_{n,k} = \sqrt{\frac{n}{2}} (a_{n,k} + a_{n,k}^+) = \frac{1}{\sqrt{2}} (\alpha_{n,k} + \alpha_{-n,k}).$$

The nonvanishing commutators are given by

$$[x_{n,k}, p_{m,l}] = iG_{kl} \delta_{n,m}, \quad [y_k, p_l] = iG_{kl}.$$  (3.136)

We should note that the metric appearing in eqs. (3.136) is $G_{ij}$, instead
of $g_{ij}$. So it can be seen that if we employ the variables with the lowered
space–time indices $y_k, p_k, x_{n,k}$ and $p_{n,k}$, the metric used in the expression
of the overlaps is $G^{ij}$ not $g^{ij}$.

The continuity condition (3.125) is universal for any background, and
the mode expansion of the momenta $P_i(\sigma)$'s is of the same form as in the
Neumann case, thus the continuity conditions for the momenta in terms of
the modes $p_{n,i}$ are identical with those in the Neumann case. Also, since
$p_{n,i}$'s mutually commute, it is natural to find a solution of the continuity
condition, assuming the following form for the overlap vertices:

$$|\hat{V}_N\rangle_{X_{1-N}} = \exp \left[ \frac{i}{4\pi \alpha'} \theta^{ij} \sum_{r,s=1}^{N} p_{n,i}^{(r)} Z_{nm}^{rs} p_{m,j}^{(s)} \right] |V_N\rangle_{X_{1-N}},$$  (3.137)

where $|\hat{V}_N\rangle_{X_{1-N}}$ and $|V_N\rangle_{X_{1-N}}$ are the overlaps in the background corresponding to the world sheet actions (3.126) and (3.122) respectively, the explicit form of the latter is given in appendix A. Clearly the expression (3.137)
satisfies the continuity conditions for the modes of the momenta,
and the coefficients $Z_{mn}^s$ are determined so that the continuity conditions for the coordinates are satisfied [Sugino (2000)].

- $|\hat{I}X \equiv |\hat{V}_1X$

For the $N = 1$ case, we consider the identity overlap $|\hat{I}X \equiv |\hat{V}_1X$. The continuity conditions for the momenta require that

$$P_i(\sigma) + P_i(\pi - \sigma) = \frac{2}{\pi} p_i + \frac{2}{\pi \sqrt{\alpha'}} \sum_{n=2,4,6,\ldots} \cos(n\sigma) p_{n,i}$$

should vanish for $0 \leq \sigma \leq \frac{\pi}{2}$, namely,

$$p_i = 0, \quad p_{n,i} = 0 \quad (n = 2, 4, 6, \ldots) \label{3.138}$$

which is satisfied by the overlap in the Neumann case $|I\rangle$. In addition, the conditions for the coordinates are that

$$X^j(\sigma) - X^j(\pi - \sigma) = \frac{2}{\pi} \theta^{j,k}(\sigma - \frac{\pi}{2}) p_k + 4\sqrt{\alpha'} \sum_{n=1,3,5,\ldots} G^{jk} \cos(n\sigma) x_{n,k} + 4\sqrt{\alpha'} \sum_{n=2,4,6,\ldots} \frac{1}{2\pi \alpha'} \theta^{j,k} \sin(n\sigma) \frac{1}{n} p_{n,k}, \label{3.139}$$

should vanish for $0 \leq \sigma \leq \frac{\pi}{2}$. The first and third lines in the r. h. s. can be put to zero by using (3.138). So what we have to consider is the remaining condition $x_{n,k} = 0$ for $n = 1, 3, 5, \ldots$, which however is nothing but the continuity condition for the coordinates in the Neumann case. It can be understood from the point that the second line in (3.139) does not depend on $\theta^{ij}$. Thus it turns out that the continuity conditions in the case of the $B-$field turned on are satisfied by the identity overlap made in the Neumann case. The solution is [Sugino (2000)]

$$|\hat{I}X \equiv |I\rangle X = \exp \left[ -\frac{1}{2} G^{ij} \sum_{n=0}^{\infty} (-1)^n a_{n,i}^\dagger a_{n,j}^\dagger \right] |0\rangle, \label{3.140}$$

where also the zero modes $y_i$ and $p_i$ are written by using the creation and annihilation operators $a_{0,i}^\dagger$ and $a_{0,i}$ as

$$y_i = \frac{1}{2} \sqrt{2\alpha'}(a_{0,i} - a_{0,i}^\dagger), \quad p_i = \frac{1}{\sqrt{2\alpha'}}(a_{0,i} + a_{0,i}^\dagger).$$
For the \( N = 2 \) case, we are to do the same argument as in the \( N = 1 \) case. The continuity conditions mean that

\[
P_i^{(1)}(\sigma) + P_i^{(2)}(\pi - \sigma) = \frac{1}{\pi}(P_i^{(1)} + P_i^{(2)}) + \frac{1}{\pi \sqrt{\alpha'}} \sum_{n=1}^{\infty} \cos(n\sigma)(p_{n,i}^{(1)} + (-1)^n p_{n,i}^{(2)}),
\]

\[
X^{(1)}(\sigma) - X^{(2)}(\pi - \sigma) = G^{jk}(y_k^{(1)} - y_k^{(2)}) + \frac{1}{\pi} \theta^{jk}(\sigma - \frac{\pi}{2})(p_k^{(1)} + p_k^{(2)}) + 2\sqrt{\alpha'} \sum_{n=1}^{\infty} \left[ G^{jk} \cos(n\sigma)(x_n^{(1)} - (-1)^n x_n^{(2)}) \right.
\]

\[
+ \left. \frac{1}{2\pi \alpha'} \theta^{jk} \sin(n\sigma) \frac{1}{n} p_{n,k}^{(1)} + (-1)^n p_{n,k}^{(2)} \right]
\]

should be zero for \( 0 \leq \sigma \leq \pi \). It turns out again that the conditions for the modes are identical with those in the Neumann case:

\[
p_i^{(1)} + p_i^{(2)} = 0, \quad p_{n,i}^{(1)} + (-1)^n p_{n,i}^{(2)} = 0,
\]

\[
y_i^{(1)} - y_i^{(2)} = 0, \quad x_n^{(1)} - (-1)^n x_n^{(2)} = 0,
\]

for \( n \geq 1 \). Thus we have the solution

\[
|\hat{V}_2^2\rangle^{X}_{12} = |V_2^2\rangle^{X}_{12} = \exp \left[ -G^{ij} \sum_{n=0}^{\infty} (-1)^n a_{n,i}^{(1)} a_{n,j}^{(2)} \right] |0\rangle_{12}. \quad (3.141)
\]

• |\( \hat{V}_4 \rangle^{X}_{1234}\)

We find a solution of the continuity conditions \( 3.125 \) in the \( N = 4 \) case assuming the form

\[
|\hat{V}_4\rangle^{X}_{1234} = \exp \left[ \frac{i}{4\pi \alpha} \theta^{ij} \sum_{r,s=1}^{4} p_{n,r}^{(r)} G_{ns}^{(s)} P_{m,j} \right] |V_4\rangle^{X}_{1234}. \quad (3.142)
\]

When considering the continuity conditions, it is convenient to employ the
For the momentum variables we also define the \( Z \) variables following mode expansions \( \bar{P} \) instead of putting the subscript \( t \) where \( \bar{P} \) instead of \( P \) the same combinations of \( P \) of \( t \) which can be seen from the point that the conditions (3.144) for the sectors \( \leq 4 \) for \( 0 = 2 \) and 4 are identical with the Neumann case using those variables, the continuity conditions are written as

\[
Q_1^j(\sigma) = \frac{1}{2} [i X^{(1)}j(\sigma) - X^{(2)}j(\sigma) - i X^{(3)}j(\sigma) + X^{(4)}j(\sigma)] \equiv Q^j(\sigma),
\]

\[
Q_2^j(\sigma) = \frac{1}{2} [-X^{(1)}j(\sigma) + X^{(2)}j(\sigma) - X^{(3)}j(\sigma) + X^{(4)}j(\sigma)],
\]

\[
Q_3^j(\sigma) = \frac{1}{2} [-i X^{(1)}j(\sigma) - X^{(2)}j(\sigma) + i X^{(3)}j(\sigma) + X^{(4)}j(\sigma)] \equiv Q^j(\sigma),
\]

\[
Q_4^j(\sigma) = \frac{1}{2} [X^{(1)}j(\sigma) + X^{(2)}j(\sigma) + X^{(3)}j(\sigma) + X^{(4)}j(\sigma)].
\]

For the momentum variables we also define the \( Z_4 \)-Fourier transformed variables \( P_{t,i}(\sigma) \equiv P_t(\sigma) \), \( P_{2,i}(\sigma) \), \( P_{3,i}(\sigma) \equiv P_t(\sigma) \) and \( P_{4,i}(\sigma) \) by the same combinations of \( P^{(t)}_i(\sigma) \)'s as the above. These variables have the following mode expansions

\[
P_{t,i}(\sigma) = \frac{1}{\pi \sqrt{2 \alpha'}} P_{t,0,i} + \frac{1}{\pi \sqrt{\alpha'}} \sum_{n=1}^{\infty} \cos(n \sigma) P_{t,n,i},
\]

\[
Q_1^j(\sigma) = G^{jk} \sqrt{2 \alpha'} Q_{t,0,k} + \frac{1}{\pi} \theta^{jk}(\sigma - \frac{\pi}{2}) \frac{1}{\sqrt{2 \alpha'}} P_{2,0,k}
\]

\[+ \sqrt{2 \alpha'} \sum_{n=1}^{\infty} \left[ G^{jk} \cos(n \sigma) Q_{t,n,k} + \frac{1}{2 \pi \alpha'} \theta^{jk} \sin(n \sigma) \frac{1}{n} P_{t,n,k} \right], \tag{3.143}
\]

where \( t = 1, 2, 3, 4 \). From now on, we frequently omit the subscript \( t \) for the \( t = 1 \) case, and at the same time we employ the expression with a bar instead of putting the subscript \( t \) for the \( t = 3 \) case.

Using those variables, the continuity conditions are written as

\[
Q_1^j(\sigma) - Q_1^j(\pi - \sigma) = 0, \quad P_{4,i}(\sigma) + P_{4,i}(\pi - \sigma) = 0,
\]

\[
Q_2^j(\sigma) + Q_2^j(\pi - \sigma) = 0, \quad P_{2,i}(\sigma) - P_{2,i}(\pi - \sigma) = 0,
\]

\[
Q_3^j(\sigma) - i Q_3^j(\pi - \sigma) = 0, \quad P_{3,i}(\sigma) + i P_{3,i}(\pi - \sigma) = 0,
\]

\[
Q_4^j(\sigma) + i Q_4^j(\pi - \sigma) = 0, \quad P_{4,i}(\sigma) - i P_{4,i}(\pi - \sigma) = 0 \tag{3.144}
\]

for \( 0 \leq \sigma \leq \frac{\pi}{2} \). In terms of the modes, the conditions for the sectors of \( t = 2 \) and 4 are identical with the Neumann case

\[
(1 - C)|Q_{4,k}^j| \bar{V}_4^X = (1 + C)|P_{4,k}^j| \bar{V}_4^X = 0,
\]

\[
(1 + C)|Q_{2,k}^j| \bar{V}_4^X = (1 - C)|P_{2,k}^j| \bar{V}_4^X = 0,
\]

which can be seen from the point that the conditions (3.144) for the sectors of \( t = 2 \) and 4 lead the same relations between the modes as those without
the terms containing $\theta^{jk}$. Here we adopted the vector notation for the modes

$$|Q_{t,k}⟩ = \begin{bmatrix} Q_{t,0,k} \\ Q_{t,1,k} \\ \vdots \end{bmatrix}, \quad |P_{t,k}⟩ = \begin{bmatrix} P_{t,0,k} \\ P_{t,1,k} \\ \vdots \end{bmatrix},$$

and $C$ is a matrix such that $(C)_{nm} = (-1)^n \delta_{nm}$ ($n, m \geq 0$). Thus there is needed no correction containing $\theta^{ij}$ for the sectors of $t = 2$ and 4, so it is natural to assume the form of the phase factor in (3.142) as

$$\frac{1}{2} \theta^{ij} \sum_{r,s=1}^{4} (p_{r}^{(i)}|Z^{rs}|p_{s}^{(j)}) = \theta^{ij} (P_{i}|Z|P_{j})$$

with $Z$ being anti–Hermitian.

Next let us consider the conditions for the sectors of $t = 1$ and 3. We rewrite the mode expansions of $Q_{j}(σ)$ and $\bar{Q}_{j}(σ)$ as [Sugino (2000)]

$$Q_{j}(σ) = G^{jk} (\sqrt{2\alpha′} Q_{0,k} + 2\sqrt{\alpha′} \sum_{n=1}^{∞} \cos(nσ) Q_{n,k}) + \theta^{jk} \left[ \int_{\pi/2}^{σ} dσ′ P_{i}(σ′) + \frac{1}{\pi \sqrt{\alpha′}} \sum_{n=1,3,5,\ldots}^{∞} \frac{1}{n} (-1)^{(n-1)/2} P_{n,k} \right]$$

$$\equiv \theta^{jk} \int_{\pi/2}^{σ} dσ′ P_{i}(σ′) + \Delta Q^j(σ), \quad (3.146)$$

$$\bar{Q}_{j}(σ) = G^{jk} (\sqrt{2\alpha′} \bar{Q}_{0,k} + 2\sqrt{\alpha′} \sum_{n=1}^{∞} \cos(nσ) \bar{Q}_{n,k}) + \theta^{jk} \left[ \int_{\pi/2}^{σ} dσ′ \bar{P}_{i}(σ′) + \frac{1}{\pi \sqrt{\alpha′}} \sum_{n=1,3,5,\ldots}^{∞} \frac{1}{n} (-1)^{(n-1)/2} \bar{P}_{n,k} \right]$$

$$\equiv \theta^{jk} \int_{\pi/2}^{σ} dσ′ \bar{P}_{i}(σ′) + \Delta \bar{Q}^j(σ). \quad (3.147)$$

Using the conditions for $P_{i}(σ)$ and $\bar{P}_{i}(σ)$ in (3.144), we can reduce the conditions for $Q^j(σ)$ and $\bar{Q}^j(σ)$ to those for $\Delta Q^j(σ)$ and $\Delta \bar{Q}^j(σ)$:

$$\Delta Q^j(σ) = \begin{cases} i\Delta Q^j(\pi - σ) & (0 ≤ σ ≤ \frac{π}{2}) \\ -i\Delta Q^j(π - σ) & (\frac{π}{2} ≤ σ ≤ π) \end{cases},$$

$$\Delta \bar{Q}^j(σ) = \begin{cases} -i\Delta \bar{Q}^j(π - σ) & (0 ≤ σ ≤ \frac{π}{2}) \\ i\Delta \bar{Q}^j(π - σ) & (\frac{π}{2} ≤ σ ≤ π) \end{cases}.$$
These formulas are translated to the relations between the modes via the Fourier transformation. The result is expressed in the vector notation as
\[(1 - X)|Q_i⟩|\hat{V}_4⟩^X = (1 + X)|\overline{Q}_i⟩|\hat{V}_4⟩^X = 0, \tag{3.148}\]
where the vectors $|Q_i⟩$ and $|\overline{Q}_i⟩$ stand for
\[|Q_i⟩ = \begin{bmatrix} Q_{0,i} + \frac{i}{\sqrt{2}}G_{ik}\sum_{n=0}^{\infty} X_{0n}P_{n,j} \\ Q_{1,i} \\ Q_{2,i} \\ \vdots \end{bmatrix}, \]
\[|\overline{Q}_i⟩ = \begin{bmatrix} \bar{Q}_{0,i} + \frac{i}{\sqrt{2}}G_{ik}\sum_{n=0}^{\infty} X_{0n}\bar{P}_{n,j} \\ \bar{Q}_{1,i} \\ \bar{Q}_{2,i} \\ \vdots \end{bmatrix}. \]

In (3.148), passing the vectors through the phase factor of $|\hat{V}_4⟩^X$ and using the continuity conditions in the Neumann case
\[(1 + X)|P_i⟩|V_4⟩^X = (1 - X)|\bar{P}_i⟩|V_4⟩^X = 0, \tag{3.149}\]
we get the equations, which the coefficients $Z_{nm}$’s should satisfy,
\[[(1 - X)_{m0}\sum_{n=0}^{\infty}(Z_{0n} + i\frac{\pi}{2}\bar{X}_{0n})P_{n,j} + \sum_{n=1}^{\infty}(1 - X)_{mn}\sum_{n'=0}^{\infty} Z_{nm}P_{n',j}]|V_4⟩^X = 0 \]
\[[(1 + X)_{m0}\sum_{n=0}^{\infty}(Z_{0n} - i\frac{\pi}{2}X_{0n})\bar{P}_{n,j} + \sum_{n=1}^{\infty}(1 + X)_{mn}\sum_{n'=0}^{\infty} Z_{nm}\bar{P}_{n',j}]|V_4⟩^X = 0 \]
for $m \geq 0$. Now all our remaining task is to solve these equations. It is easy to see that a solution of them is given by [Sugino (2000)]
\[Z_{mn} = -i\frac{\pi}{2}(1 - X)_{mn} + i\beta\frac{\pi}{2}C_{mn}, \quad (m, n \geq 0, \text{ except for } m = n = 0), \]
\[Z_{00} = i\beta\frac{\pi}{2}, \]
if we pay attention to (3.149). Here $\beta$ is an unknown real constant, which is not fixed by the continuity conditions alone. This ambiguity of the solution comes from the property of the matrix $X$: $XC = -CX$. However it will become clear that the term containing the constant $\beta$ does not contribute to the vertex $|\hat{V}_4⟩^X$. 

---

**Feynman Path Integrals**

191
Therefore, we have the expression of the phase (3.145)

\[
\theta^{ij}(P_i|Z|\bar{P}_j) = \theta^{ij}[\pi \sum_{n=0}^{\infty} (-1)^n P_n,i \bar{P}_n,j + i \beta \sum_{m,n=0}^{\infty} P_{m,i}(1 - X)_{m,n} \bar{P}_{n,j}].
\]

Then recalling (3.149) again, the last term in the r. h. s. can be discarded. Also we can rewrite the term containing \(\beta - \theta^{ij}4\alpha'\beta\)

\[
\frac{\theta^{ij}}{4\alpha'}(P_i|C|\bar{P}_j) = \frac{\theta^{ij}}{4\alpha'}(P_i|X^T CX|\bar{P}_j) = \frac{\theta^{ij}}{4\alpha'}(P_i|C|\bar{P}_j),
\]

on \(|V_4\rangle^X\). The above formula means that the term containing \(\beta\) can be set to zero on \(|V_4\rangle^X\). After all, the form of the 4-string vertex becomes

\[
|\hat{V}_4\rangle_{1234}^X = \exp \left[-\frac{\theta^{ij}}{4\alpha'} P_{0,i} \bar{P}_{0,j}\right] |V_4\rangle_{1234}^X.
\]

Note that the phase factor has the cyclic symmetric form

\[
-\frac{\theta^{ij}}{4\alpha'} P_{0,i} \bar{P}_{0,j} = i \frac{\theta^{ij}}{8\alpha' P_{0,i} P_{0,j}} + p_{0,i} P_{0,j} + p_{0,i} P_{0,j} + p_{0,i} P_{0,j},
\]

which is a property the vertices should have.\(^{19}\)

• \(|\hat{V}_3\rangle_{123}^X\)

We can get the 3-string overlap in the similar manner as in the 4-string case. First, we introduce the \(Z_3\)-Fourier transformed variables

\[
Q_1^j(\sigma) = \frac{1}{\sqrt{3}}[e^{iX^{(1)}}(\sigma) + \bar{e}^{iX^{(2)}}(\sigma) + X^{(3)}(\sigma)] \equiv Q^j(\sigma),
\]

\[
Q_2^j(\sigma) = \frac{1}{\sqrt{3}}[e^{iX^{(1)}}(\sigma) + \bar{e}^{iX^{(2)}}(\sigma) + X^{(3)}(\sigma)] \equiv \bar{Q}^j(\sigma),
\]

\[
Q_3^j(\sigma) = \frac{1}{\sqrt{3}}[X^{(1)}(\sigma) + X^{(2)}(\sigma) + X^{(3)}(\sigma)],
\]

where \(e \equiv e^{i2\pi/3}\), \(\bar{e} \equiv e^{-i2\pi/3}\). The momenta \(P_{1,i}(\sigma)(\equiv P_{1}(\sigma))\), \(P_{2,i}(\sigma)(\equiv \bar{P}_{1}(\sigma))\) and \(P_{3,i}(\sigma)\) are defined in the same way. The mode expansions

\(^{19}\)Here the momentum \(p_{0,i}^{(r)}\) is given by \(p_{0,i}^{(r)} = \sqrt{2\alpha'}p_i^{(r)}\).
take the same form as those in (3.143). In these variables, the continuity conditions require
\[ Q^j(\sigma) - eQ^j(\pi - \sigma) = 0, \quad P^j(\sigma) + eP^j(\pi - \sigma) = 0, \]
\[ Q^j_3(\sigma) - \bar{Q}^j_3(\pi - \sigma) = 0, \quad P^j_3(\sigma) + \bar{P}^j_3(\pi - \sigma) = 0, \]
for \( 0 \leq \sigma \leq \frac{\pi}{2} \). The conditions imposed to the modes with respect to the \( t = 3 \) component are identical with those in the Neumann case
\[ (1 + C)|P_{3,i}||\bar{V}_3|^X = (1 - C)|Q_{3,i}||\bar{V}_3|^X = 0. \]
Thus the \( t = 3 \) component does not couple with \( \theta^j \), so we can find the solution by determining the single anti–Hermitian matrix \( Z \) in the phase factor whose form is assumed as \([\text{Sugino (2000)}]\)
\[ \frac{1}{2} \theta^{ij} \sum_{r,s=1}^3 (p^{(r)}_i|Z^{rs}|p^{(s)}_j) = \theta^{ij}(P_i|Z|P_j). \quad (3.150) \]

For the sectors of \( t = 1 \) and \( 2 \), the same argument goes on as in the 4-string case. \( Q^j(\sigma) \) and \( \bar{Q}^j(\sigma) \) have the mode expansions same as in eqs. (3.146) and (3.147). The conditions we have to consider are
\[ \Delta Q^j(\sigma) = \begin{cases} e\Delta Q^j(\pi - \sigma), & (0 \leq \sigma \leq \frac{\pi}{2}) \\ e\Delta Q^j(\pi - \sigma), & (\frac{\pi}{2} \leq \sigma \leq \pi), \end{cases} \]
\[ \Delta \bar{Q}^j(\sigma) = \begin{cases} e\Delta \bar{Q}^j(\pi - \sigma), & (0 \leq \sigma \leq \frac{\pi}{2}) \\ e\Delta \bar{Q}^j(\pi - \sigma), & (\frac{\pi}{2} \leq \sigma \leq \pi), \end{cases} \]
which are rewritten as the relations between the modes
\[ (1 - Y)|Q_i||\bar{V}_3|^X = (1 - Y^T)|\bar{Q}_i||\bar{V}_3|^X = 0. \quad (3.151) \]
Recalling the conditions in the Neumann case
\[ (1 + Y)|P_i||\bar{V}_3|^X = (1 + Y^T)|\bar{P}_i||\bar{V}_3|^X = 0, \quad (3.152) \]
\[ (1 - Y)|Q_i||\bar{V}_3|^X = (1 - Y^T)|\bar{Q}_i||\bar{V}_3|^X = 0, \]
we end up with the following equations
\[ [(1 - Y)_{mn} \sum_{n=0}^{\infty} \left( Z_{0n} + \frac{\pi}{2} \bar{X}_{0n} \right) P_{n,j} + \sum_{n=1}^{\infty} (1 - Y)_{mn} \sum_{n'=0}^{\infty} Z_{n'n'} P_{n',j'}]|V_3|^X = 0, \]
\[ [(1 - Y^T)_{mn} \sum_{n=0}^{\infty} \left( Z_{0n} - \frac{\pi}{2} \bar{X}_{0n} \right) \bar{P}_{n,j} + \sum_{n=1}^{\infty} (1 - Y^T)_{mn} \sum_{n'=0}^{\infty} Z_{n'n'} \bar{P}_{n',j'}]|V_3|^X = 0 \]
for $m \geq 0$. It can be easily found out that the expression

$$Z_{mn} = -i\frac{\pi}{\sqrt{3}}(1 + Y^T)_{mn} \quad (m, n \geq 0, \text{ except for } m = n = 0),$$

$$Z_{00} = 0,$$

satisfies the above equations. It should be noted that in this case, because of $CYC = \tilde{Y} \neq -Y$, it does not contain any unknown constant differently from the 4-string case.

Owing to the condition (3.152) we can write the phase factor only in terms of the zero-modes. Finally we have (Sugino (2000))

$$|\hat{V}_3\rangle^{X}_{123} = \exp\left[-i\frac{\theta^{ij}}{4\sqrt{3}\alpha'}P_{0,i}\tilde{P}_{0,j}\right]|V_3\rangle^{X}_{123} = \exp\left[i\frac{\theta^{ij}}{12\alpha'}(p_{0,i}^{(1)}p_{0,j}^{(2)} + p_{0,i}^{(2)}p_{0,j}^{(3)} + p_{0,i}^{(3)}p_{0,j}^{(1)})\right]|V_3\rangle^{X}_{123}.$$  (3.153)

It is not clear whether the solutions we have obtained here are unique or not. However we can show that the phase factors are consistent with the relations between the overlaps which they should satisfy,

$$3\langle\hat{I}|\hat{V}_3\rangle_{123} = |\hat{V}_2\rangle_{12}, \quad 4\langle\hat{I}|\hat{V}_4\rangle_{1234} = |\hat{V}_3\rangle_{123}, \quad 3\langle\hat{V}_2||\hat{V}_3\rangle_{123}|\hat{V}_4\rangle_{3456} = |\hat{V}_4\rangle_{1256}.$$  

by using the momentum conservation on the vertices $(p_{1}^{(1)} + \cdots + p_{N}^{(N)}|V_N\rangle^{X}_{1,\ldots,N} = 0$. Furthermore we can see that the phase factors successfully reproduce the Moyal product structures of the correlators among vertex operators obtained in the perturbative approach to open string theory in the constant $B-$field background (Seiberg and Witten (1999)). These facts convince us that the solutions obtained here are physically meaningful.

3.3.6.4 Transformation of String Fields

In the previous section, we have explicitly constructed the overlap vertices in the operator formulation under the constant $B-$field background. Then we have obtained the vertices with a new noncommutative structure of the Moyal type originating from the constant $B-$field, in addition to the ordinary product $*$ of string fields. Denoting the product with the new
structure by $\star$, the action of the string field theory is written as

$$S_B = \frac{1}{G_s} \int \left( \frac{1}{2} \psi \star Q \psi + \frac{1}{3} \psi \star \psi \star \psi \right) = \frac{1}{G_s} \left( \frac{1}{2} \langle \hat{V}_2 || \psi \rangle_1 Q | \psi \rangle_2 + \frac{1}{3} \langle \hat{V}_3 || \psi \rangle_1 | \psi \rangle_2 | \psi \rangle_3 \right), \quad (3.154)$$

where the BRST charge $Q$ is constructed from the world sheet action (3.126). The theory (3.154) gives the noncommutative $U(1)$ Yang–Mills theory in the low energy region in the same sense as Witten’s open string field theory in the case of the Neumann boundary condition leads to the ordinary $U(1)$ Yang–Mills theory in the low energy limit.

In [Seiberg and Witten (1999)] the authors argued that open string theory in the constant $B$–field background leads to either commutative or noncommutative Yang–Mills theories, corresponding to the different regularization scheme (the so–called Pauli–Villars regularization or the point–splitting regularization) in the world sheet formulation. They discussed a map between the gauge fields in the commutative and noncommutative Yang–Mills theories. In string field theory perspective, there also should be a certain transformation (hopefully simpler than the Yang–Mills case) from the string field $\psi$ in (3.154) to a string field in a new string field theory which leads to the commutative Yang–Mills theory in the low energy limit.

Here we get the new string field theory by finding a unitary transformation which absorbs the noncommutative structure of the Moyal type in the product $\star$ into a redefinition of the string fields. There are used the two vertices $|\hat{V}_2 \rangle$ and $|\hat{V}_3 \rangle$ in the action (3.154). Recall that the 2–string vertex is in the same form as in the Neumann case and has no Moyal type noncommutative structure. First, we consider the phase factor of the 3–string vertex which multiplies in front of $|\hat{V}_3 \rangle$ (see (3.153)). Making use of the continuity conditions

$$P_{0,i} = -2 \sum_{n=1}^{\infty} Y_{0n} P_{n,i}, \quad \bar{P}_{0,i} = -2 \sum_{n=1}^{\infty} \bar{Y}_{0n} \bar{P}_{n,i}, \quad (3.155)$$

$^{20}$It can be explicitly seen by repeating a similar calculation as that carried out in Dearnaley (1990).
it can be rewritten as \cite{Sugino (2000)}

\[
-\frac{\theta^{ij}}{4\sqrt{3}\alpha'} P_{0,i} P_{0,j} = \frac{\theta^{ij}}{4\sqrt{3}\alpha'} \sum_{n=1}^{\infty} (P_{0,i} \tilde{Y}_0 P_{n,j} + P_{n,i} Y_0 \tilde{P}_{0,j})
\]

\[
= -\frac{\theta^{ij}}{24\alpha'} \sum_{n=1}^{\infty} X_{0n} \left[ (-p_{0,i}^{(2)} - p_{0,i}^{(3)}) + 2p_{0,i}^{(1)} p_{n,j}^{(1)} \right]
\]

\[
+ (-p_{0,i}^{(3)} - p_{0,i}^{(1)}) + 2p_{0,i}^{(2)} p_{n,j}^{(2)} + (-p_{0,i}^{(1)} - p_{0,i}^{(2)}) + 2p_{0,i}^{(2)} p_{n,j}^{(3)} \right]
\]

\[
= -\frac{\theta^{ij}}{8\alpha'} \sum_{n=1}^{\infty} X_{0n} p_{0,i}^{(r)} p_{0,j}^{(r)}
\]

where we used the property of the matrix \( Y \): \( Y_0 = -\tilde{Y}_0 = \sqrt{\frac{3}{2}} X_0 \) for \( n \geq 1 \) and the momentum conservation on \( |V_3\rangle \): \( p_{0,i}^{(1)} + p_{0,i}^{(2)} + p_{0,i}^{(3)} = 0 \). We manage to represent the phase factor of the Moyal type as a form factorized into the product of the unitary operators

\[
U_r = \exp \left( \frac{\theta^{ij}}{8\alpha'} \sum_{n=1,3,5,...} X_{0n} p_{0,i}^{(r)} p_{0,j}^{(r)} \right).
\] (3.156)

Note that the unitary operator acts to a single string field. So the Moyal type noncommutativity can be absorbed by the unitary rotation of the string field

\[
12\langle \hat{V}_3|\psi_1\psi_2\psi_3 \rangle = 123 \langle V_3|U_1 U_2 U_3|\psi_1\psi_2\psi_3 \rangle = 123 \langle V_3|\tilde{\psi}_1\tilde{\psi}_2\tilde{\psi}_3 \rangle.
\] (3.157)

with \( |\tilde{\psi}_r\rangle = U_r |\psi_r\rangle \). It should be remarked that this manipulation has been succeeded owing to the factorized expression of the phase factor, which originates from the continuity conditions relating the zero-modes to the nonzero-modes \( \tilde{X} \). It is a characteristic feature of string field theory that can not be found in any local field theories.

Next let us see the kinetic term. In doing so, it is judicious to write the kinetic term as follows:

\[
12\langle \hat{V}_2|\psi_1\psi_2 \rangle = 123 \langle V_3|Q_L |I_2|\psi_3 \rangle + \langle \psi_2 Q_L |I_3 \rangle
\] (3.158)

where \( Q_L \) is defined by integrating the BRST current \( j_{BRST}(\sigma) \) with respect to \( \sigma \) over the left half region

\[
Q_L = \int_0^{\pi/2} d\sigma j_{BRST}(\sigma).
\]
Equation (3.158) is also represented by the product $\star$ as
\[ \psi \star (Q \psi) = \psi \star [(Q_L I) \star \psi + \psi \star (Q_L I)]. \] (3.159)
Here, $I$ stands for the identity element with respect to the $\star$ product, carrying the ghost number $-\frac{3}{2}$, which corresponds to $|I\rangle$ in the operator formulation. As is discussed by Horowitz et al. (1986), in order to show the relation (3.159) we need the formulas
\[ Q_R I = -Q_L I, \quad (Q_R \psi) \star \xi = -(-1)^{n_\psi} \psi \star (Q_L \xi) \] (3.160)
for arbitrary string fields $\psi$ and $\xi$, where $Q_R$ is the integrated BRST current over the right half region of $\sigma$. $n_\psi$ stands for the ghost number of the string field $\psi$ minus $\frac{1}{2}$, and takes an integer value. The first formula means that the identity element is a physical quantity, also the second does the conservation of the BRST charge. By using these formulas, the first term in the bracket in r. h. s. of (3.159) becomes
\[ (Q_L I) \star \psi = -(Q_R I) \star \psi = I \star (Q_L \psi) = Q_L \psi. \]
Also, it turns out that the second term is equal to $Q_R \psi$. Combining these, we can see that (3.159) holds.

Further, we should remark that because the BRST current does not contain the center of mass coordinate $x_j$, it commute with the momentum $p_i$. From the continuity condition $p_i |I\rangle = 0$, it can be seen that $p_i Q_L |I\rangle = 0$. Expanding the exponential in the expression of the unitary operator (3.156) and passing the momentum $p_{0,i}$ to the right, we get
\[ U Q_L |I\rangle = Q_L |I\rangle. \] (3.161)
Now we can write down the result of the kinetic term. As a result of the same manipulation as in eq. (3.157) and the use of eq. (3.161), we have\footnote{Strictly speaking, in general this formula holds in the case that both of the string fields $|\psi\rangle$ and $|\tilde{\psi}\rangle$ belong to the Fock space which consists of states excited by finite number of creation operators. This point is subtle for giving a proof. However, for the infinitesimal $\theta$ case, by keeping arbitrary finite order terms in the expanded form of the exponential of $U$, we can make the situation of both $|\psi\rangle$ and $|\tilde{\psi}\rangle$ being inside the Fock space, and thus clearly eq. (3.162) holds. From this fact, it is plausible to expect that eq. (3.162) is correct in the finite $\theta$ case.}
\[ 12 (\tilde{V}_2 |\psi\rangle_1 (Q |\psi\rangle_2) = 12 \langle \tilde{V}_2 |\psi\rangle_1 (Q_L |I\rangle_2 |\psi\rangle_3 + |\psi\rangle_2 Q_L |I\rangle_3) \\
= 12 \langle \tilde{V}_2 |\psi\rangle_1 (Q_L |I\rangle_2 |\tilde{\psi}\rangle_3 + |\tilde{\psi}\rangle_2 Q_L |I\rangle_3) = 12 \langle \tilde{V}_2 |\psi\rangle_1 (|\tilde{\psi}\rangle_2). \] (3.162)
Here we have a comment [Sugino (2000)]. If we considered the kinetic term itself without using (3.158), what would be going on? Let us see this. From the continuity conditions for \(|\hat{V}_2\rangle_{12} = |V_2\rangle_{12}\):

\[
p^{(1)}_{0,i} + p^{(2)}_{0,i} = 0, \quad p^{(1)}_{n,i} + (-1)^n p^{(2)}_{n,i} = 0 \quad (n = 1, 2, \cdots),
\]

it could be shown that the 2–string overlap is invariant under the unitary rotation

\[
U_1 U_2 |V_2\rangle_{12} = |V_2\rangle_{12}.
\]

So we would find the expression for the kinetic term after the rotation

\[
1_2 \langle V_2|\tilde{Q}|\psi\rangle_1 = 1_2 \langle V_2|\tilde{Q}|\tilde{\psi}\rangle_2,
\]

where \(\tilde{Q}\) is the BRST charge similarity transformed by \(U\)

\[
\tilde{Q} = UQU^\dagger.
\] (3.163)

However, after some computations of the r. h. s. of (3.163), we could see that \(\tilde{Q}\) has divergent term proportional to

\[
\sum_{n=1,3,5,\cdots} 1
\]

and thus it is not well–defined. It seems that this procedure is not correct and needs some suitable regularization, which preserves the conformal symmetry\(^\text{22}\). It is considered that the use of eq. (3.158) gives that kind of regularization, which will be justified at the end of the next section.

Therefore, the string field theory action (3.154) with the Moyal type noncommutativity added to the ordinary noncommutativity is equivalently

\[^{22}\text{That divergence comes from the mid–point singularity of the string coordinates transformed by } U. \text{ In fact, after some calculations, we have}
\]

\[
U X^j(\sigma) U^\dagger = X^j(\sigma) - \frac{\theta^{jk}}{4\sqrt{2\alpha'}} \sum_{n=1,3,5,\cdots} X_{n0p_{00},k} - \frac{\theta^{jk}}{4} p_k \text{ sgn} \left(\sigma - \frac{\pi}{2}\right).
\] (3.164)

The last term leads to the mid-point singularity in the energy–momentum tensor and the BRST charge \(Q\). It seems that the use of (3.158) corresponds to taking the point splitting regularization with respect to the mid–point. Because of the discontinuity of the last term in (3.164), it is considered that the transformed string coordinates have no longer a good picture as a string. It could be understood from the point that the transformation \(U\) drives states around a perturbative vacuum to those around highly non–perturbative one like coherent states.
rewritten as the one with the ordinary noncommutativity alone \cite{Sugino (2000)}:

\[
S_B = \frac{1}{G_s} \int \left( \frac{1}{2} \bar{\psi} * Q \psi + \frac{1}{3} \bar{\psi} * \bar{\psi} * \psi \right)
= \frac{1}{G_s} \left( \frac{1}{2} \langle V_2 | \langle \psi \rangle_1 Q | \bar{\psi} \rangle_2 + \frac{1}{3} \langle V_3 | \langle \psi \rangle_1 | \bar{\psi} \rangle_2 | \bar{\psi} \rangle_3 \right). \tag{3.165}
\]

It is noted that the BRST charge \( Q \), which is constructed from the world sheet action \eqref{W1}, has the same form as the one obtained from the action \eqref{W2} with the relation \eqref{W3}. So all the \( B \)–dependence has been stuffed into the string fields except that existing in the metric \( G_{ij} \). Furthermore, recalling that the relation between the metrics \( G_{ij} \) and \( g_{ij} \) is the same form as the T–duality inversion transformation, which was pointed out at the end of section 3, we can make the metric \( g_{ij} \) appear in the overlap vertices, instead of the metric \( G_{ij} \). To do so, we consider the following transformation for the modes:

\[
\hat{\alpha}_i^j = (E^T)^{ik}\alpha_{n,k}, \quad \hat{p}_i^j = (E^T)^{ik}p_{k}, \quad \hat{x}_i = E_{ik}x^k. \tag{3.166}
\]

By this transformation, the commutators become

\[
[\hat{\alpha}_i^j, \hat{\alpha}_m^k] = ng^{ij}\delta_{n+m,0}, \quad [\hat{p}_i^j, \hat{x}_j] = -i\delta_j^i,
\]

and the bilinear form of the modes

\[
G_{ij} \alpha_{n,j} \alpha_{m,j} = g_{ij} \hat{\alpha}_i^j \hat{\alpha}_m^j, \quad G_{ij} p_i \alpha_{m,j} = g_{ij} \hat{p}_i^j \hat{\alpha}_m^j, \quad G_{ij} p_i p_j = g_{ij} \hat{p}_i^j \hat{p}_j^j. \tag{3.167}
\]

### 3.3.7 Topological Strings

The 2D field theories we have constructed are already very similar to string theories. However, one ingredient from string theory is missing: in string theory, the world–sheet theory does not only involve a path integral over the maps \( \phi^j \) to the target space and their fermionic partners, but also a path integral over the world–sheet metric \( h_{\alpha\beta} \). So far, we have set this metric to a fixed background value.

We have also encountered a drawback of our construction. Even though the theories we have found can give us some interesting ‘semi–topological’ information about the target spaces, one would like to be able to define general nonzero \( n \)–point functions at genus \( g \) instead of just the partition function.
function at genus one and the particular correlation functions we calculated at genus zero.

It turns out that these two remarks are intimately related. In this section we will go from topological field theory to topological string theory by introducing integrals over all metrics, and in doing so we will find interesting nonzero correlation functions at any genus (see [Vonk (2005)]).

**Coupling to Topological Gravity**

In coupling an ordinary field theory to gravity, one has to perform the following three steps.

- First of all, one rewrites the Lagrangian of the theory in a covariant way by replacing all the flat metrics by the dynamical ones, introducing covariant derivatives and multiplying the measure by a factor of $\sqrt{\text{det } h}$.
- Secondly, one introduces an Einstein–Hilbert term as the ‘kinetic’ term for the metric field, plus possibly extra terms and fields to preserve the symmetries of the original Lagrangian.
- Finally, one has to integrate the resulting theory over the space of all metrics.

Here we will not discuss the first two steps in this procedure. As we have seen in our discussion of topological field theories, the precise form of the Lagrangian only plays a comparatively minor role in determining the properties of the theory, and we can derive many results without actually considering a Lagrangian. Therefore, let us just state that it is possible to carry out the analog of the first two steps mentioned above, and construct a Lagrangian with a ‘dynamical’ metric which still possesses the topological $Q$–symmetry we have constructed. The reader who is interested in the details of this construction is referred to the paper [Witten (1990)] and to the lecture notes [Dijkgraaf et al. (1991)].

The third step, integrating over the space of all metrics, is the one we will be most interested in here. Naively, by the metric independence of our theories, integrating their partition functions over the space of all metrics, and then dividing the results by the volume of the topological ‘gauge group’, would be equivalent to multiplication by a factor of 1,

$$Z[h_0] = \frac{1}{G_{\text{top}}} \int D[h] Z[h], \quad (3.168)$$

for any arbitrary background metric $h_0$. There are several reasons why this naive reasoning might go wrong:
• There may be metric configurations which cannot be reached from a given metric by continuous changes.
• There may be anomalies in the topological symmetry at the quantum level preventing the conclusion that all gauge fixed configurations are equivalent.
• The volume of $G_{\text{top}}$ is infinite, so even if we could rigorously define a path integral the above multiplication and division would not be mathematically well-defined.

For these reasons, we should really be more careful and precisely define what we mean by the ‘integral over the space of all metrics’. Let us note the important fact that just like in ordinary string theory (and even before twisting), the 2D sigma models become conformal field theories when we include the metric in the Lagrangian. This means that we can borrow the technology from string theory to integrate over all conformally equivalent metrics. As is well known, and as we will discuss in more detail later, the conformal symmetry group is a huge group, and integrating over conformally equivalent metrics leaves only a $nD$ integral over a set of world-sheets moduli. Therefore, our strategy will be to use the analogy to ordinary string theory to first do this integral over all conformally equivalent metrics, and then perform the integral over the remaining $nD$ moduli space.

In integrating over conformally equivalent metrics, one usually has to worry about conformal anomalies. However, here a very important fact becomes our help. To understand this fact, it is useful to rewrite our twisting procedure in a somewhat different language (see [Vonk (2005)]).

Let us consider the SEM-tensor $T_{\alpha\beta}$, which is the conserved Noether current with respect to global translations on $C$. From conformal field theory, it is known that $T_{zz} = T_{\bar{z}\bar{z}} = 0$, and the fact that $T$ is a conserved current, $\partial_\alpha T^\alpha_{\beta} = 0$, means that $T_{zz} \equiv T(z)$ and $T_{\bar{z}\bar{z}} \equiv \bar{T}(\bar{z})$ are (anti-) holomorphic in $z$. One can now expand $T(z)$ in Laurent modes,

$$T(z) = \sum L_m z^{-m-2}. \quad (3.169)$$

The $L_m$ are called the Virasoro generators, and it is a well-known result from conformal field theory that in the quantum theory their commutation relations are

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{c}{12}m(m^2-1)\delta_{m+n}. $$

The number $c$ depends on the details of the theory under consideration, and it is called the central charge. When this central charge is nonzero, one
runs into a technical problem. The reason for this is that the equation of motion for the metric field reads

$$\frac{\delta S}{\delta h^{\alpha \beta}} = T_{\alpha \beta} = 0.$$ 

In conformal field theory, one imposes this equation as a constraint in the quantum theory. That is, one requires that for physical states $|\psi\rangle$,

$$L_m |\psi\rangle = 0 \quad (\text{for all } m \in \mathbb{Z}).$$

However, this is clearly incompatible with the above commutation relation unless $c = 0$. In string theory, this value for $c$ can be achieved by taking the target space of the theory to be 10D. If $c \neq 0$ the quantum theory is problematic to define, and we speak of a “conformal anomaly” [Vonk (2005)].

The whole above story repeats itself for $\bar{T}(\bar{z})$ and its modes $\bar{L}_m$. At this point there is a crucial difference between open and closed strings. On an open string, left–moving and right–moving vibrations are related in such a way that they combine into standing waves. In our complex notation, ‘left–moving’ translates into ‘$z$–dependent’ (i.e., holomorphic), and ‘right–moving’ into ‘$\bar{z}$–dependent’ (i.e., anti–holomorphic). Thus, on an open string all holomorphic quantities are related to their anti–holomorphic counterparts. In particular, $T(z)$ and $\bar{T}(\bar{z})$, and their modes $L_m$ and $\bar{L}_m$, turn out to be complex conjugates. There is therefore only one independent algebra of Virasoro generators $L_m$.

On a closed string on the other hand, which is the situation we have been studying so far, left– and right–moving waves are completely independent. This means that all holomorphic and anti–holomorphic quantities, and in particular $T(z)$ and $\bar{T}(\bar{z})$, are independent. One therefore has two sets of Virasoro generators, $L_m$ and $\bar{L}_m$.

Let us now analyze the problem of central charge in the twisted theories. To twist the theory, we have used the $U(1)$–symmetries. Any global $U(1)$–symmetry of our theory has a conserved current $J_z$. The fact that it is conserved again means that $J_z \equiv J(z)$ is holomorphic and $\bar{J}_{\bar{z}} \equiv \bar{J}(\bar{z})$ is anti–holomorphic. Once again, on an open string $J$ and $\bar{J}$ will be related, but in the closed string theory we are studying they will be independent functions. In particular, this means that we can view a global $U(1)$–symmetry as really consisting of two independent, left– and right–moving, $U(1)$–symmetries, with generators $F_L$ and $F_R$.

Note that the sum of $U(1)$–symmetries $F_Y + F_A$ only acts on objects with a $+$ index. That is, it acts purely on left–moving quantities. Simi-
larly, \( F_V - F_A \) acts purely on right–moving quantities. From our discussion above, it is therefore natural to identify these two symmetries with the two components of a single global \( U(1) \) symmetry:

\[
F_V = \frac{1}{2}(F_L + F_R), \quad F_A = \frac{1}{2}(F_L - F_R).
\]

A more detailed construction shows that this can indeed be done.

Let us expand the left–moving conserved \( U(1) \)–current into Laurent modes,

\[
J(z) = \sum J_m z^{-m-1}.
\] (3.170)

The commutation relations of these modes with one another and with the Virasoro modes can be calculated, either by writing down all of the modes in terms of the fields of the theory, or by using more abstract knowledge from the theory of superconformal symmetry algebras. In either case, one finds

\[
[L_m, L_n] = (m - n)L_{m+n} + \frac{c}{12}m(m^2 - 1)\delta_{m+n}\left[J_m, J_n\right]
\]

\[
= -nJ_{m+n}\left[J_m, J_n\right] = \frac{c}{3}m\delta_{m+n}.
\]

Note that the same central charge \( c \) appears in the \( J \)– and in the \( L \)–commutators. This turns out to be crucial.

Following the standard Noether procedure, we can now construct a conserved charge by integrating the conserved current \( J(z) \) over a space–like slice of the \( z \)–plane. In string theory, the physical time direction is the radial direction in the \( z \)–plane, so a space–like slice is just a curve around the origin. The integral is therefore calculated using the Cauchy Theorem,

\[
F_L = \oint_{z=0} J(z)\,dz = 2\pi iJ_0.
\]

In the quantum theory, it will be this operator that generates the \( U(1)_L \)–symmetry. Now recall that to twist the theory we want to introduce new Lorentz rotation generators,

\[
M_A = M - F_V = M - \frac{1}{2}(F_L + F_R), \quad M_B = M - F_A = M - \frac{1}{2}(F_L - F_R).
\]

A well–known result from string theory (see \cite{Vonk (2005)}) is that the generator of Lorentz rotations is \( M = 2\pi i(L_0 - \bar{L}_0) \). Therefore, we find
that the twisting procedure in this new language amounts to

\[ A : L_{0,A} = L_0 - \frac{1}{2} J_0, \quad \bar{L}_{0,A} = \bar{L}_0 + \frac{1}{2} \bar{J}_0, \]

\[ B : L_{0,B} = L_0 - \frac{1}{2} J_0, \quad \bar{L}_{0,B} = \bar{L}_0 - \frac{1}{2} \bar{J}_0. \]

Let us now focus on the left-moving sector; we see that for both twistings the new Lorentz rotation generator is the difference of \( L_0 \) and \( \frac{1}{2} J_0 \). The new Lorentz generator should also correspond to a conserved 2-tensor, and from (3.169) and (3.170) there is a very natural way to get such a current:

\[ \tilde{T}(z) = T(z) + \frac{1}{2} \partial J(z), \tag{3.171} \]

which clearly satisfies \( \bar{\partial} \tilde{T} = 0 \) and

\[ \tilde{L}_m = L_m - \frac{1}{2} (m + 1) J_m, \tag{3.172} \]

so in particular we find that \( \tilde{L}_0 \) can serve as \( L_{0,A} \) or \( L_{0,B} \). We should apply the same procedure (with a minus sign in the \( A \)-model case) in the right-moving sector. Equations (3.171) and (3.172) tell us how to implement the twisting procedure not only on the conserved charges, but on the whole \( N = 2 \) superconformal algebra or at least on the part consisting of the \( J- \) and \( L- \) modes, but a further investigation shows that this is the only part that changes. We have motivated, but not rigorously derived (3.171); for a complete justification the reader is referred to the original papers [Lerche et al. (1989)] and [Cecotti and Vafa (1991)].

Now, we come to the crucial point. The algebra that the new modes \( \tilde{L}_m \) satisfy can be directly calculated from (3.171), and we find

\[ [\tilde{L}_m, \tilde{L}_n] = (m - n) \tilde{L}_{m+n}. \]

That is, there is no central charge left. This means that we do not have any restriction on the dimension of the theory, and topological strings will actually be well-defined in target spaces of any dimension.

From this result, we see that we can integrate our partition function over conformally equivalent metrics without having to worry about the conformal anomaly represented by the nonzero central charge. After having integrated over this large part of the space of all metrics, it turns out that there is a \( nD \) integral left to do. In particular, it is known that one can always find a conformal transformation which in the neighborhood of a chosen point puts the metric in the form \( h_{\alpha \beta} = \eta_{\alpha \beta} \), with \( \eta \) the usual flat
metric with diagonal entries ±1. (Or, +1 in the Euclidean setting.) On the other hand, when one considers the global situation, it turns out that one cannot always enforce this gauge condition everywhere. For example, if the world-sheet is a torus, there is a left-over complex parameter $\tau$ that cannot be gauged away. The easiest way to visualize this parameter (see [Vonk (2005)]) is by drawing the resulting torus in the complex-plane and rescaling it in such a way that one of its edges runs from 0 to 1; the other edge then runs from 0 to $\tau$, see Figure 3.19. It seems intuitively clear that a conformal transformation – which should leave all angles fixed – will never deform $\tau$, and even though intuition often fails when considering conformal mappings, in this case this can indeed be proven. Thus, $\tau$ is really a modular parameter which we need to integrate over. Another fairly intuitive result is that any locally flat torus can, after a rescaling, be drawn in this form, so $\tau$ indeed is the only modulus of the torus.

![Fig. 3.19 The only modulus $\tau$ of a torus $T^2$.](image)

More generally, one can show that a Riemann surface of genus $g$ has $m_g = 3(g - 1)$ complex modular parameters. As usual, this is the virtual dimension of the moduli space. If $g > 1$, one can show that this virtual dimension equals the actual dimension. For $g = 0$, the sphere, we have a negative virtual dimension $m_g = -3$, but the actual dimension is 0: there is always a flat metric on a surface which is topologically a sphere (just consider the sphere as a plane with a point added at infinity), and after having chosen this metric there are no remaining parameters such as $\tau$ in the torus case. For $g = 1$, the virtual dimension is $m_g = 0$, but as we have seen the actual dimension is 1.

We can explain these discrepancies using the fact that, after we have used the conformal invariance to fix the metric to be flat, the sphere and the
torus have leftover symmetries. In the case of the sphere, it is well known in string theory that one can use these extra symmetries to fix the positions of three labelled points. In the case of the torus, after fixing the metric to be flat we still have rigid translations of the torus left, which we can use to fix the position of a single labelled point. To see how this leads to a difference between the virtual and the actual dimensions, let us for example consider tori with \( n \) labelled points on them. Since the virtual dimension of the moduli space of tori without labelled points is 0, the virtual dimension of the moduli space of tori with \( n \) labelled points is \( n \). One may expect that at some point (and in fact, this happens already when \( n = 1 \)), one reaches a sufficiently generic situation where the virtual dimension really is the actual dimension. However, even in this case we can fix one of the positions using the remaining conformal (translational) symmetry, so the positions of the points only represent \( n - 1 \) moduli. Hence, there must be an \( n \)th modulus of a different kind, which is exactly the shape parameter \( \tau \) that we have encountered above. In the limiting case where \( n = 0 \), this parameter survives, thus causing the difference between the virtual and the real dimension of the moduli space.

For the sphere, the reasoning is somewhat more formal: we analogously expect to have three ‘extra’ moduli when \( n = 0 \). In fact, three extra parameters are present, but they do not show up as moduli. They must be viewed as the three parameters which need to be added to the problem to find a 0D moduli space.

Since the cases \( g = 0, 1 \) are thus somewhat special, let us begin by studying the theory on a Riemann surface with \( g > 1 \). To arrive at the topological string correlation functions, after gauge fixing we have to integrate over the remaining moduli space of complex dimension \( 3(g - 1) \). To do this, we need to fix a measure on this moduli space. That is, given a set of \( 6(g - 1) \) tangent vectors to the moduli space, we want to produce a number which represents the size of the volume element spanned by these vectors, see Figure 3.20. We should do this in a way which is invariant under coordinate redefinitions of both the moduli space and the world-sheet. Is there a ‘natural’ way to do this?

To answer this question, let us first ask how we can describe the tangent vectors to the moduli space (see \([\text{Vonk (2005)}]\)). In two dimensions, conformal transformations are equivalent to holomorphic transformations: \( z \mapsto f(z) \). It thus seems natural to assume that the moduli space we have left labels different complex structures on \( \Sigma \), and indeed this can be shown to be the case. Therefore, a tangent vector to the moduli space is an
Fig. 3.20 A measure on the moduli space $M$ assigns a number to every set of three tangent vectors. This number is interpreted as the volume of the element spanned by these vectors.

An infinitesimal change of complex structure, and these changes can be parameterized by holomorphic 1–forms with anti–holomorphic vector indices,

$$dz \mapsto dz + \epsilon \mu^z_i(z) d\bar{z}.$$ 

The dimension counting above tells us that there are $3(g - 1)$ independent $(\mu_i)_z$, plus their $3(g - 1)$ complex conjugates which change $d\bar{z}$. So the tangent space is spanned by these $\mu_i(z, \bar{z}), \bar{\mu}_i(z, \bar{z})$. How do we get a number out of a set of these objects? Since $\mu_i$ has a $z$ and a $\bar{z}$ index, it seems natural to integrate it over $\Sigma$. However, the $z$–index is an upper index, so we need to lower it first with some tensor with two $z$–indices. It turns out that a good choice is to use the $Q$–partner $G_{zz}$ of the SEM–tensor component $T_{zz}$, and thus to define the integration over moduli space as

$$\int_M \prod_{i=1}^{3g-3} \left( dm^i dm^i \int_{\Sigma} G_{zz}(\mu_i)_z^z \int_{\Sigma} G_{zz}(\bar{\mu}_i)^z \right).$$  

(3.173)

Note that by construction, this integral is also invariant under a change of basis of the moduli space. There are several reasons why using $G_{zz}$ is a natural choice. First of all, this choice is analogous to what one does in bosonic string theory. There, one integrates over the moduli space using exactly the same formula, but with $G$ replaced by the conformal ghost $b$. This ghost is the BRST–partner of the SEM–tensor in exactly the same way as $G$ is the $Q$–partner of $T$. Secondly, one can make the not unrelated observation that since $\{Q, G\} = T$, we can still use the standard arguments to show independence of the theory of the parameters in a Lagrangian of the form $L = \{Q, V\}$. The only difference is that now we also have to commute $Q$ through $G$ to make it act on the vacuum, but since $T_{\alpha\beta}$ itself is the derivative of the action with respect to the metric $h_{\alpha\beta}$, the terms we get
in this way amount to integrating a total derivative over the moduli space. Therefore, apart from possible boundary terms these contributions vanish. Note that this reasoning also gives us an argument for using $G_{zz}$ instead of $T_{zz}$ (which is more or less the only other reasonable option) in (3.173): if we had chosen $T_{zz}$ then all path integrals would have been over total derivatives on the moduli space, and apart from boundary contributions the whole theory would have become trivial.

If we consider the vector and axial charges of the full path integral measure, including the new path integral over the world–sheet metric $h$, we find a surprising result. Since the world–sheet metric does not transform under $R$–symmetry, naively one might expect that its measure does not either. However, this is clearly not correct since one should also take into account the explicit $G$–insertions in (3.173) that do transform under $R$–symmetry. From the $N = 2$ superconformal algebra (or, more down–to–earth, from expressing the operators in terms of the fields), it follows that the product of $G$ and $\bar{G}$ has vector charge zero and axial charge 2. Therefore, the total vector charge of the measure remains zero, and the axial charge gets an extra contribution of $6(g - 1)$, so we find a total axial $R$–charge of $6(g - 1) - 2m(g - 1)$. From this, we see that the case of complex target space dimension 3 is very special: here, the axial charge of the measure vanishes for any $g$, and hence the partition function is nonzero at every genus. If $m > 3$ and $g > 1$, the total axial charge of the measure is negative, and we have seen that we cannot cancel such a charge with local operators. Therefore, for these theories only the partition function at $g = 1$ and a specific set of correlation functions at genus zero give nonzero results. Moreover, for $m = 2$ and $m = 1$, the results can be shown to be trivial by other arguments. Therefore, a Calabi–Yau threefold is by far the most interesting target space for a topological string theory. It is a ‘happy coincidence’ (see [Vonk (2005)]) that this is exactly the dimension we are most interested in from the string theory perspective.

Finally, let us come back to the special cases of genus 0 and 1. At genus zero, the Riemann surface has a single point as its moduli space, so there are no extra integrals or $G$–insertions to worry about. Therefore, we can copy the topological field theory result saying that we have to introduce local operators with total degree $(m, m)$ in the theory. The only remnant of the fact that we are integrating over metrics is that we should also somehow fix the remaining three symmetries of the sphere. The most straightforward way to do this is to consider 3–point functions with insertions on three labelled points. As a gauge choice, we can then for example require these points
to be at the points 0, 1 and $\infty$ in the compactified complex–plane. For example, in the $A$–model on a Calabi–Yau threefold, the 3–point function of three operators corresponding to $(1, 1)$–forms would thus give a nonzero result.

In the case of the torus, we have seen that there is one ‘unexpected’ modular parameter over which we have to integrate. This means we have to insert one $G$– and one $\bar{G}$–operator in the measure, which spoils the absence of the axial anomaly we had for $g = 1$ in the topological field theory case. However, we also must fix the one remaining translational symmetry, which we can do by inserting a local operator at a labelled point. Thus, we can restore the axial $R$–charge to its zero value by choosing this to be an operator of degree $(1, 1)$.

Summarizing, we have found that in topological string theory on a target Calabi–Yau 3–fold, we have a non–vanishing 3–point function of total degree $(3,3)$ at genus zero; a non–vanishing 1–point function of degree $(1,1)$ at genus one, and a non–vanishing partition (‘zero–point’) function at all genera $g > 1$.

**Nonlocal Operators**

In one respect, what we have achieved is great progress: we can now for any genus define a nonzero partition function (or for low genus a correlation function) of the topological string theory. On the other hand, we would also like to define correlation functions of an arbitrary number of operators at these genera. As we have seen, the insertion of extra local operators in the correlation functions is not possible, since any such insertion will spoil our carefully constructed absence of $R$–symmetry anomalies. Therefore, we have to introduce nonlocal operators.

There is one class of nonlocal operators which immediately becomes mind. Before we saw, using the descent equations, that for every local operator we can define a corresponding 1–form and a 2–form operator. If we check the axial and vector charges of these operators, we find that if we start with an operator of degree $(1, 1)$, the 2–form operator we end up with actually has vanishing axial and vector charges. This has two important consequences. First of all, we can add the integral of this operator to our action [Vonk (2005)],

$$S[t] = S_0 + \int t^a \mathcal{O}^{(2)}_a,$$
without spoiling the axial and vector symmetry of the theory. Secondly, we can insert the integrated operator into correlation functions,

\[ \langle \int O^{(2)}_1 \cdots \int O^{(2)}_n \rangle \]

and still get a nonzero result by the vanishing of the axial and vector charges. These two statements are related: one obtains such correlators by differentiating \( S[t] \) with respect to the appropriate \( t \)'s, and then setting all \( t^a = 0 \).

A few remarks are in place here. First of all, recall that the integration over the insertion points of the operators can be viewed as part of the integration over the moduli space of Riemann surfaces, where now we label a certain number of points on the Riemann surface. From this point of view, the \( g = 0,1 \) cases fit naturally into the same framework. We could unite the descendant fields into a world–sheet super–field,

\[ \Phi_a = O^{(0)}_a + O^{(1)}_{a\alpha} \theta^\alpha + O^{(2)}_{a\alpha\beta} \theta^\alpha \bar{\theta}^\beta \]

where we formally replaced each \( dz \) and \( d\bar{z} \) by corresponding fermionic coordinates \( \theta^z \) and \( \bar{\theta}^\bar{z} \). Now, one can write the above correlators as integrals over \( n \) copies of this super–space,

\[ \int \prod_{s=1}^n d^2 z_s d^2 \theta_s \langle \Phi_{a_1}(z_1, \theta_1) \cdots \Phi_{a_n}(z_n, \theta_n) \rangle \]

The integration prescription at genus 0 and 1 tells us to fix 3 and 1 points respectively, so we need to remove this number of super–space integrals. Then, integrating over the other super–space coordinates, the genus 0 correlators indeed become

\[ \langle O^{(0)}_{a_1} O^{(0)}_{a_2} O^{(0)}_{a_3} \int O^{(2)}_{a_4} \cdots \int O^{(2)}_{a_n} \rangle \]

From this prescription we note that these expressions are symmetric in the exchange of all \( a_i \) and \( a_j \). In particular, this means that the genus zero 3–point functions at arbitrary \( t \),

\[ c_{abc}[t] = \langle O^{(0)}_a O^{(0)}_b O^{(0)}_c \rangle[t] \]

have symmetric derivatives:

\[ \frac{\partial c_{abc}}{\partial t^a} = \frac{\partial c_{abcd}}{\partial t^d} \]
and similarly with permuted indices. These equations can be viewed as integrability conditions, and using the Poincaré lemma we see that they imply that

\[ c_{ijk}[t] = \frac{\partial Z_0[t]}{\partial t^i \partial t^j \partial t^k}. \]

for some function \( Z_0[t] \). Following the general philosophy that \( n \)-point functions are \( n \)th derivatives of the \( t \)-dependent partition function, we see that \( Z_0[t] \) can be naturally thought of as the partition function at genus zero. Similarly, the partition function at genus 1 can be defined by integrating up the one-point functions once.

The quantities we have calculated above should be semi-topological invariants, meaning that they only depend on ‘half’ of the moduli (either the Kähler ones or the complex structure ones) of the target space. For example, in the \( A \)-model we find the Gromov–Witten invariants. In the \( B \)-model, it turns out that \( F_0[t] = \ln Z_0[t] \) is actually a quantity we already knew: it is the prepotential of the Calabi–Yau manifold. A discussion of why this is the case can be found in the paper [Bershadsky et al. (1994)]. The higher genus partition functions can be thought of as ‘quantum corrections’ to the prepotential.

Finally, there is a type of operator we have not discussed at all so far. Recall that in the topological string theory, the metric itself is now a dynamical field. We could not include the metric in our physical operators, since this would spoil the topological invariance. However, the metric is part of a \( Q \)-multiplet, and the highest field in this multiplet is a scalar field which is usually labelled \( \phi \). (It should not be confused with the fields \( \phi^i \).) We can get more correlation functions by inserting operators \( \phi^k \) and the operators related to them by the descent equations into the correlation functions. These operators are called ‘gravitational descendants’. Even the case where the power is \( k = 0 \) is nontrivial; it does not insert any operator, but it does label a certain point, and hence changes the moduli space one integrates over. This operator is called the ‘puncture operator’.

All of this seems to lead to an enormous amount of semi-topological target space invariants that can be calculated, but there are many recursion relations between the several correlators. This is similar to how we showed before that all correlators for the cohomological field theories follow from the 2- and 3-point functions on the sphere. Here, it turns out that the set of all correlators has a structure which is related to the theory of integrable hierarchies. Unfortunately, a discussion of this is outside the scope of both
these lectures and the author’s current knowledge.

The Holomorphic Anomaly

We have now defined the partition function and correlation functions of topological string theory, but even though the expressions we obtained are much simpler than the path integrals for ordinary quantum field or string theories, it would still be very hard to explicitly calculate them. Fortunately, it turns out that the $t$–dependent partition and correlation functions are actually ‘nearly holomorphic’ in $t$, and this is a great aid in exactly calculating these quantities.

Let us make this ‘near holomorphy’ more precise. As we have seen, calculating correlation functions of primary operators in topological string theories amounts to taking $t$–derivatives of the corresponding perturbed partition function $Z[t]$ and consequently setting $t = 0$. Recall that $Z[t]$ is defined through adding terms to the action of the form

$$ t^a \int \Sigma \mathcal{O}_a^{(2)}, $$

(3.174)

Let us for definiteness consider the $B$–twisted model. We want to show that the above term is $Q_B$–exact. For simplicity, we assume that $\mathcal{O}_a^{(2)}$ is a bosonic operator, but what we are about to say can by inserting a few signs straightforwardly be generalized to the fermionic case. From the descent equations we studied above, we know that

$$ (\mathcal{O}_a^{(2)})_{+-+-} = -\{G_+, [G_-, \mathcal{O}_a^{(0)}]\}, $$

(3.175)

where $G_+$ is the charge corresponding to the current $G_{zz}$, and $G_-$ the one corresponding to $G_{zz}$. We can in fact express $G_\pm$ in terms of the $N = (2, 2)$ supercharges $Q$. So, according to [Vonk (2005)], we have

$$ H = 2\pi i(L_0 + \bar{L}_0) = \frac{1}{2}\{Q_+, \bar{Q}_+\} - \frac{1}{2}\{Q_-, \bar{Q}_-\}P $$

$$ = 2\pi i(L_0 - \bar{L}_0) = \frac{1}{2}\{Q_+, \bar{Q}_+\} + \frac{1}{2}\{Q_-, \bar{Q}_-\}. $$

Thus, we find that the left– and right–moving SEM charges satisfy

$$ T_+ = 2\pi iL_0 = \frac{1}{2}\{Q_+, \bar{Q}_+\} T_- = 2\pi i\bar{L}_0 = -\frac{1}{2}\{Q_-, \bar{Q}_-\}. $$
To find $G$ in the $B$–model, we should write these charges as commutators with respect to $Q_B = \bar{Q}_+ + \bar{Q}_-$, which gives

$$T_+ = \frac{1}{2} \{Q_B, Q_+ \}, \quad T_- = -\frac{1}{2} \{Q_B, Q_- \},$$

so we arrive at the conclusion that for the $B$–model,

$$G_+ = \frac{1}{2} Q_+ G_- = -\frac{1}{2} Q_-.$$

Now, we can rewrite (3.175) as

$$\left( O^{(2)}_a \right) = -\{G_+, [G_-, O^{(0)}_a]\} = \frac{1}{4} \{Q_+, [Q_-, O^{(0)}_a]\}$$

(3.176)

$$= \frac{1}{8} \{Q_B, [(Q_- - Q_+), O^{(0)}_a]\},$$

which proves our claim that $O^{(2)}_a$ is $Q_B$–exact.

An $N = (2, 2)$ sigma model with a real action does, apart from the term (3.174), also contain a term

$$t^a \int_\Sigma O^{(2)}_a,$$

(3.177)

where $t^a$ is the complex conjugate of $t^a$. It is not immediately clear that $O^{(2)}_a$ is a physical operator: we have seen that physical operators in the $B$–model correspond to forms that are $\bar{\partial}$–closed, but the complex conjugate of such a form is $\partial$–closed. However, by taking the complex conjugate of (3.176), we see that

$$\left( O^{(2)}_a \right) = \frac{1}{8} \{Q_B, [(Q_- - Q_+), O^{(0)}_a]\},$$

so not only is the operator $Q_B$–closed, it is even $Q_B$–exact. This means that we can add terms of the form (3.177) to the action, and taking $t^a$–derivatives inserts $Q_B$–exact terms in the correlation functions. Naively, we would expect this to give a zero result, so all the physical quantities seem to be $t$–independent, and thus holomorphic in $t$. We will see in a moment that this naive expectation turns out to be almost right, but not quite.

However, before doing so, let us comment briefly on the generalization of the above argument in the case of the $A$–model. It seems that a straightforward generalization of the argument fails, since $Q_A$ is its own complex conjugate, and the complex conjugate of the de Rham operator is also the same operator. However, note that the $N = (2, 2)$–theory has a
different kind of ‘conjugation symmetry’: we can exchange the two super-
symmetries, or in other words, exchange $\theta^+$ with $\bar\theta^+$ and $\theta^-$ with $\bar\theta^-$. This
exchanges $Q_A$ with an operator which we might denote as $Q_A \equiv Q_+ + Q_-$. Using the above argument, we then find that the physical operators $\mathcal{O}_a^{(2)}$ are $Q_A$–exact, and that their conjugates in the new sense are $Q_A$–exact.
We can now add these conjugates to the action with parameters $t^a$, and we again naively find independence of these parameters. In this case it is less natural to choose $t^a$ and $t^\bar{a}$ to be complex conjugates, but we are free to choose this particular ‘background point’ and study how the theory behaves if we then vary $t^a$ and $t^\bar{a}$ independently.

Now, let us see how the naive argument showing independence of the theory of $t^\bar{a}$ fails. In fact, the argument above would certainly hold for topological field theories. However, in topological string theories (see [Vonk (2005)]), we have to worry about the insertions in the path integral of
\[
G \cdot \mu_i \equiv \int d^2 z G_{zz} (\mu_i) \bar{z},
\]
and their complex conjugates, when commuting the $Q_B$ towards the vacuum and making sure it gives a zero answer. Indeed, the $Q_B$–commutator of the above factor is not zero, but it gives
\[
\{Q_B, G \cdot \mu_i \} = T \cdot \mu_i.
\]
Now recall that $T_{\alpha\beta} = \partial h_{\alpha\beta} S$. We did not give a very precise definition of $\mu_i$ above, but we know that it parameterizes the change in the metric under an infinitesimal change of the coordinates $m_i$ on the moduli space. One can make this intuition precise, and then finds the following ‘chain rule’: $T \cdot \mu_i = \partial m_i S$. Inserting this into the partition function, we find that
\[
\frac{\partial F_g}{\partial t^a} = 
\int_{M_g} \prod_{i=1}^{3g-3} d m_i d \bar{m}_i \sum_{j,k} \frac{\partial^2}{\partial m_j \partial \bar{m}_k} \left\langle \left( \prod_{l \neq j} \int \mu_l \cdot G \right) \left( \prod_{l \neq k} \int \bar{\mu}_l \cdot \bar{G} \right) \int \mathcal{O}_a^{(2)} \right\rangle,
\]
where $F_g = \ln Z_g$ is the free energy at genus $g$, and the reason $F_g$ appears in the above equation instead of $Z_g$ is, as usual in quantum field theory, that the expectation values in the r.h.s. are normalized such that $\langle 1 \rangle = 1$, and so the l.h.s. should be normalized accordingly and equal $Z^{-1}_g \partial_h Z_g = \partial_h F_g$ [Vonk (2005)].
Thus, as we have claimed before, we are integrating a total derivative over the moduli space of genus $g$ surfaces. If the moduli space did not have a boundary, this would indeed give zero, but in fact the moduli space does have a boundary. It consists of the moduli which make the genus $g$ surface degenerate. This can happen in two ways: an internal cycle of the genus $g$ surface can be pinched, leaving a single surface of genus $g - 1$, as in Figure 3.21(a), or the surface can split up into two surfaces of genus $g_1$ and $g_2 = g - g_1$, as depicted in Figure 3.21(b). By carefully considering the boundary contributions to the integral for these two types of boundaries, it was shown in [Bershadsky et al. (1994)] that

$$\frac{\partial F_g}{\partial t^a} = \frac{1}{2} c_{\bar{a} \bar{b} \bar{c}} e^{2K} G^{\bar{d} \bar{e}} G^{\bar{e} \bar{d}} \left( D_{\bar{d}} D_{\bar{e}} F_{g-1} + \sum_{r=1}^{g-1} D_{\bar{d}} F_r D_{\bar{e}} F_{g-r} \right),$$

where $G$ is the so-called Zamolodchikov metric on the space parameterized by the coupling constants $t^a, \bar{t}^\bar{a}$; $K$ is its Kähler potential, and the $D_a$ are covariant derivatives on this space. The coefficients $c_{\bar{a} \bar{b} \bar{c}}$ are the 3-point functions on the sphere of the operators $\bar{O}_{a(0)}$. We will not derive the above formula in detail, but the reader should notice that the contributions from the two types of boundary are quite clear.

![Fig. 3.21](image.png)

**Fig. 3.21** At the boundary of the moduli space of genus $g$ surfaces, the surfaces degenerate because certain cycles are pinched. This either lowers the genus of the surface (a) or breaks the surface into two lower genus ones (b) (see text for explanation).

Using this formula, one can inductively determine the $t^a$ dependence on the partition functions if the holomorphic $t^\alpha$–dependence is known. Holomorphic functions on complex spaces (or more generally holomorphic sections of complex vector bundles) are quite rare: usually, there is only a $nD$ space of such functions. The same turns out to hold for our topological string partition functions: even though they are not quite holomorphic, their anti–holomorphic behavior is determined by the holomorphic
dependence on the coordinates, and as a result there is a finite number of coefficients which determines them.

Thus, just from the above structure and without doing any path integrals, one can already determine the topological string partition functions up to a finite number of constants. This leads to a feasible program for completely determining the topological string partition function for a given target space and at given genus. From the holomorphic anomaly equation, one first has to find the general form of the partition function. Then, all one has left to do is to fix the unknown constants. Here, the fact that in the $A$–model the partition function counts a number of points becomes our help: by requiring that the $A$–model partition functions are integral, one can often fix the unknown constants and completely determine the $t$–dependent partition function. In practice, the procedure is still quite elaborate, so we will not describe any examples here, but several have been worked out in detail in the literature. Once again, the pioneering work for this can be found in the paper [Bershadsky et al. (1994)].

3.3.8 Geometrical Transitions

Conifolds

Recall that a conifold is a generalization of the notion of a manifold. Unlike manifolds, a conifold can (or, should) contain conical singularities i.e., points whose neighborhood looks like a cone with a certain base. The base is usually a 5D manifold.

In string theory, a conifold transition represents such an evolution of the Calabi–Yau manifold in which its fabric rips and repairs itself, yet with mild and acceptable physical consequences in the context of string theory. However, the tears involved are more severe than those in an ‘weaker’ flop transition (see [Greene (2000)]). The geometrically singular conifolds were shown to lead to completely smooth physics of strings. The divergences are ‘smeared out’ by D3–branes wrapped on the shrinking 3–sphere $S^3$, as originally pointed out by A. Strominger, who, together with D. Morrison and B. Greene have also found that the topology near the conifold singularity can undergo a topological phase–transition (see subsection 3.4.5). It is believed that nearly all Calabi–Yau manifolds can be connected via these ‘critical transitions’.

More precisely, the conifold is the simplest example of a non–compact
Calabi–Yau 3–fold: it is the set of solutions to the equation
\[ x_1 x_2 - x_3 x_4 = 0 \]
in \( \mathbb{C}^4 \). The resulting manifold is a cone, meaning in this case that any real multiple of a solution to this equation is again a solution. The point \((0,0,0,0)\) is the ‘tip’ of this cone, and it is a singular point of the solution space. Note that by writing
\[ x_1 = z_1 + iz_2, \quad x_2 = z_1 - iz_2, \quad x_3 = z_3 + iz_4, \quad x_4 = -z_3 + iz_4, \]
where the \( z_i \) are still complex numbers, one can also write the equation as
\[ z_1^2 + z_2^2 + z_3^2 + z_4^2 = 0. \]
Writing each \( z_i \) as \( a_i + ib_i \), with \( a_i \) and \( b_i \) real, we get the two equations
\[ |a|^2 - |b|^2 = 0, \quad a \cdot b = 0. \] (3.178)
Here \( a \cdot b = \sum_i a_i b_i \) and \( |a|^2 = a \cdot a \). Since the geometry is a cone, let us focus on a ‘slice’ of this cone given by
\[ |a|^2 + |b|^2 = 2r^2, \]
for some \( r \in \mathbb{R} \). On this slice, the first equation in (3.178) becomes
\[ |a|^2 = r^2, \] (3.179)
which is the equation defining a 3–sphere \( S^3 \) of radius \( r \). The same holds for \( b \), so both \( a \) and \( b \) lie on 3–spheres. However, we also have to take the second equation in (3.178) into account. Let us suppose that we fix an \( a \) satisfying (3.179). Then \( b \) has to lie on a 3–sphere, but also on the plane through the origin defined by \( a \cdot b = 0 \). That is, \( b \) lies on a 2–sphere. This holds for every \( a \), so the slice we are considering is a fibration of 2–spheres over the 3–sphere. With a little more work, one can show that this fibration is trivial, so the conifold is a cone over \( S^2 \times S^3 \).

Since the conifold is a singular geometry, we would like to find geometries which approximate it, but which are non–singular. There are two interesting ways in which this can be done. The simplest way is to replace the defining equation by
\[ x_1 x_2 - x_3 x_4 = \mu^2. \] (3.180)
From the two equations constraining \( a \) and \( b \), we now see that \( |a|^2 \geq \mu^2 \). In other words, the parameter \( r \) should be at least \( \mu \). At \( r = \mu \), the
Quantum Leap

...sphere still has finite radius $\mu$, but the $b$–sphere shrinks to zero size. This geometry is called the deformed conifold. Even though this is not clear from the picture, from the equation (3.180) one can straightforwardly show that it is nonsingular. One can also show that it is topologically equivalent to the cotangent bundle on the 3–sphere, $T^*S^3$. Here, the $S^3$ on which the cotangent bundle is defined is exactly the $S^3$ at the ‘tip’ of the deformed conifold.

The second way to change the conifold geometry arises from studying the two equations

$$x_1A + x_3B = 0, \quad x_4A + x_2B = 0.$$  \hspace{1cm} (3.181)

Here, we require $A$ and $B$ to be homogeneous complex coordinates on a $\mathbb{C}P^1$, i.e.,

$$(A, B) \neq (0, 0), \quad (A, B) \sim (\lambda A, \lambda B)$$

where $\lambda$ is any nonzero complex number. If one of the $x_i$ is nonzero, say $x_1$, one can solve for $A$ or $B$, e.g., $A = -\frac{x_3B}{x_1}$, and insert this in the other equation to get

$$x_1x_2 - x_3x_4 = 0$$

which is the conifold equation. However, if all $x_i$ are zero, any $A$ and $B$ solve the system of equations (3.181). In other words, we have constructed a geometry which away from the former singularity is completely the same as the conifold, but the singularity itself is replaced by a $\mathbb{C}P^1$, which topologically is the same as an $S^2$. From the defining equations one can again show that the resulting geometry is nonsingular, so we have now replaced our conifold geometry by the so–called resolved conifold.

Topological D–branes

Since topological string theories are in many ways similar to an ordinary (bosonic) string theories, one natural question which arises is: are there also open topological strings which can end on D–branes? To answer the above question rigorously, we would have to study boundary conditions on world–sheets with boundaries which preserve the $Q$–symmetry.

In the $A$–model, one can only construct 3D–branes wrapping so–called ‘Lagrangian’ submanifolds of $M$. Here, ‘Lagrangian’ means that the Kähler form $\omega$ vanishes on this submanifold. In the $B$–model, one can construct...
D–branes of any even dimension, as long as these branes wrap holomorphic submanifolds of $M$.

Just like in ordinary string theory, when we consider open topological strings ending on a D–brane, there should be a field theory on the brane world–volume describing the low–energy physics of the open strings. Moreover, since we are studying topological theories, one may expect such a theory to inherit the property that it only depends on a restricted amount of data of the manifolds involved. A key example is the case of the $A$–model on the deformed conifold, $M = T^*S^3$, where we wrap $N$D–branes on the $S^3$ in the base. (One can show that this is indeed a Lagrangian submanifold.)

In ordinary string theory, the world–volume theory on $N$D–branes has a $U(N)$ gauge symmetry, so putting the ingredients together we can make the guess that the world–volume theory is a 3D topological field theory with $U(N)$ gauge symmetry. There is really only one candidate for such a theory: the Chern–Simons gauge theory. Recall that it consists of a single $U(N)$ gauge field, and has the action

$$S = \frac{k}{4\pi} \int_{S^3} \text{Tr} \left( A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right).$$

Before the invention of D–branes, E. Witten showed that this is indeed the theory one gets. In fact, he showed even more: this theory actually describes the full topological string–field theory on the D–branes, even without going to a low–energy limit [Witten (1995)].

Let us briefly outline the argument that gives this result. In his paper, Witten derived the open string–field theory action for the open $A$–model topological string; it reads

$$S = \int \text{Tr} \left( A * Q_A A + \frac{2}{3} A * A * A \right).$$

The form of this action is very similar to Chern–Simons theory, but its interpretation is completely different: $A$ is a string–field (a wave function on the space of all maps from an open string to the space–time manifold), $Q_A$ is the topological symmetry generator, which has a natural action on the string–field, and $*$ is a certain noncommutative product. Witten shows that the topological properties of the theory imply that only the constant maps contribute, so $A$ becomes a field on $M$ – and since open strings can only end on D–branes, it actually becomes a field on $S^3$. Moreover, recall that $Q_A$ can be interpreted as a de Rham differential. Using these observations and the precise definition of the star product one can indeed show that the
string–field theory action reduces to Chern–Simons theory on $S^3$.

### 3.3.9 Topological Strings and Black Hole Attractors

Topological string theory is naturally related to black hole dynamics (see subsection 4.3.3 below). Namely, critical string theory compactified on Calabi–Yau manifolds has played a central role in both the mathematical and physical development of modern string theory. The physical relevance of the data provided by the topological string $\hat{c} = 6$ (of A and B types) has been that it computes $F^-$–type terms in the corresponding four dimensional theory [Bershadsky et al. (1994); Antoniadis et al. (1994)]. These higher–derivative $F^-$–type terms for Type II superstring on a Calabi–Yau manifold are of the general form

$$
\int d^4x d^4\theta (W_{ab} W^{ab})^{g} F_g(X^\Lambda), \tag{3.183}
$$

where $W_{ab}$ is the graviphoton super–field of the $N = 2$ super–gravity and $X^\Lambda$ are the vector multiplet fields. The lowest component of $W$ is $F$ the graviphoton field strength and the highest one is the Riemann tensor. The lowest components of $X^\Lambda$ are the complex scalars parameterizing Calabi–Yau moduli and their highest components are the associated $U(1)$ vector–fields. These terms contribute to multiple graviphoton–graviton scattering. (3.183) includes (after $\theta$ integrations) an $R^2 F^{2g - 2}$ term. The topological string partition function $Z_{\text{top}}$ represents the canonical ensemble for multi–particle spinning five dimensional black holes [Breckenridge et al. (1997); Katz et al. (1999)]. Recently, [Ooguri et al. (2004)] proposed a simple and direct relationship between the second–quantized topological string partition function $Z_{\text{top}}$ and black hole partition function $Z_{BH}$ in four dimensions of the form

$$
Z_{BH}(p^\Lambda, \phi^\Lambda) = |Z_{\text{top}}(X^\Lambda)|^2, \quad \text{where} \quad X^\Lambda = p^\Lambda + \frac{i}{\pi} \phi^\Lambda
$$

in a certain Kähler gauge. The l.h.s. here is evaluated as a function of integer magnetic charges $p^\Lambda$ and continuous electric potentials $\phi^\Lambda$, which are conjugate to integer electric charges $q_\Lambda$. The r.h.s. is the holomorphic square of the partition function for a gas of topological strings on a Calabi–Yau whose moduli are those associated to the charges/potentials $(p^\Lambda, \phi^\Lambda)$ via the attractor equations [Ooguri et al. (2004)]. Both sides of (3.184) are defined in a perturbation expansion in $1/Q$, where $Q$ is the graviphoton
The charge carried by the black hole. The non–perturbative completion of either side of (3.184) might in principle be defined as the partition function of the holographic CFT dual to the black hole, as in [Strominger and Vafa (1996)]. Then we have the triple equality,

\[ Z_{\text{CFT}} = Z_{\text{BH}} = |Z_{\text{top}}|^2. \]

The existence of fundamental connection between 4D black holes and the topological string might have been anticipated from the following observation. Calabi–Yau spaces have two types of moduli: Kähler and complex structure. The world–sheet twisting which produces the A (B) model topological string from the critical superstring eliminates all dependence on the complex structure (Kähler) moduli at the perturbative level. Hence the perturbative topological string depends on only half the moduli. Black hole entropy on the other hand, insofar as it is an intrinsic property of the black hole, cannot depend on any externally specified moduli. What happens at leading order is that the moduli in vector multiplets are driven to attractor values at the horizon which depend only on the black hole charges and not on their asymptotically specified values. Hypermultiplet vevs on the other hand are not fixed by an attractor mechanism but simply drop out of the entropy formula. It is natural to assume this is valid to all orders in a $1/Q$ expansion. Hence the perturbative topological string and the large black hole partition functions depend on only half the Calabi–Yau moduli. It would be surprising if string theory produced two functions on the same space that were not simply related. Indeed [Ooguri et al. (2004)] argued that they were simply related as in (3.184).

**Supergravity Area–Entropy Formula**

Recall that a well–known hypothesis by J. Bekenstein and S. Hawking states that the entropy of a black hole is proportional to the area of its horizon (see [Hawking and Israel (1979)]). This area is a function of the black hole mass, or in the extremal case, of its charges. Here we review the leading semiclassical area–entropy formula for a general $N = 2$, $d = 4$ extremal black hole characterized by magnetic and electric charges $(p^A, q_A)$, recently reviewed in [Ooguri et al. (2004)]. The asymptotic values of the moduli in vector multiplets, parameterized by complex projective coordinates $X^\Lambda, (\Lambda = 0, 1, \ldots, n_V)$ in the black hole solution, are arbitrary. These

\[ 23 \] The string coupling $g_s$ is in a hypermultiplet and decouples from the computation.
moduli couple to the electromagnetic fields and accordingly vary as a function of the radius. At the horizon they approach an attractor point whose location in the moduli space depends only on the charges. The locations of these attractor points can be found by looking for supersymmetric solutions with constant moduli. They are determined by the attractor equations,

\begin{align}
    p^\Lambda &= \text{Re}[CX^\Lambda], \\
    q_\Lambda &= \text{Re}[CF_{0\Lambda}],
\end{align}

where $F_{0\Lambda} = \partial F_0 / \partial X^\Lambda$ are the holomorphic periods, and the subscript 0 distinguishes these from the string loop corrected periods to appear in the next subsection. Both $(p^\Lambda, q_\Lambda)$ and $(X^\Lambda, F_{0\Lambda})$ transform as vectors under the $Sp(2n + 2; Z)$ duality group.

The $(2n_v + 2)$ real equations (3.184) determine the $(n_v + 2)$ complex quantities $(C, X^\Lambda)$ up to Kähler transformations, which act as

\begin{align}
    K &\to K - f(X) - \tilde{f}(\bar{X}), \\
    X^\Lambda &\to e^f X^\Lambda, \\
    F_0 &\to e^{2f} F_0, \\
    C &\to e^{-f} C,
\end{align}

where the Kähler potential $K$ is given by

\[ e^{-K} = i(\bar{X}^\Lambda F_{0\Lambda} - X^\Lambda \bar{F}_{0\Lambda}). \]

We could at this point set $C = 1$ and fix the Kähler gauge but later we shall find other gauges useful. It is easy to see that (as required) the charges $(p^\Lambda, q_\Lambda)$ determined by the attractor equations (3.184) are invariant under Kähler transformations. Given the horizon attractor values of the moduli determined by (3.184), the Bekenstein–Hawking entropy $S_{BH}$ may be written as

\[ S_{BH} = \frac{1}{4} \text{Area} = \pi |Q|^2, \]

where $Q = Q_m + iQ_e$ is a complex combination of the magnetic and electric graviphoton charges and

\[ |Q|^2 = \frac{i}{2} (q_\Lambda \bar{C} X^\Lambda - p^\Lambda \bar{C} F_{0\Lambda}) = \frac{C\bar{C}}{4} e^{-K}. \]

The normalization of $Q$ here is chosen so that $|Q|$ equals the radius of the two sphere at the horizon.

It is useful to rephrase the above results in the context of type IIB superstrings in terms of geometry of Calabi–Yau. In this case the attractor equations fix the complex geometry of the Calabi–Yau. The electric/magnetic charges correlate with three cycles of Calabi–Yau. Choosing a symplectic
basis for the three cycles gives a choice of the splitting to electric and magnetic charges. Let $A_\Lambda$ denote a basis for the electric three cycles, $B^\Sigma$ the dual basis for the magnetic charges and $\Omega$ the holomorphic 3–form at the attractor point. $\Omega$ is fixed up to an overall multiplication by a complex number $\Omega \rightarrow \lambda \Omega$. There is a unique choice of $\lambda$ such that the resulting $\Omega$ has the property that

\[ p^\Lambda = \int_{A_\Lambda} \text{Re } \Omega = \text{Re}[CX^\Lambda], \quad q_\Lambda = \int_{B^\Lambda} \text{Re } \Omega = \text{Re}[CF_0^\Lambda], \]

where $\text{Re } \Omega = \frac{1}{2}(\Omega + \overline{\Omega})$.

In terms of this choice, the black hole entropy can be written as

\[ S_{BH} = \frac{\pi}{4} \int_{CY} \Omega \wedge \overline{\Omega}. \]

**Higher–Order Corrections**

$F$–term corrections to the action are encoded in a string loop corrected holomorphic prepotential

\[ F(X^\Lambda, W^2) = \sum_{h=0}^{\infty} F_h(X^\Lambda)W^{2h}, \tag{3.185} \]

where $F_h$ can be computed by topological string amplitudes (as we review in the next section) and $W^2$ involves the square of the anti–self dual graviphoton field strength. This obeys the homogeneity equation

\[ X^\Lambda \partial_\Lambda F(X^\Lambda, W^2) + W \partial_W F(X^\Lambda, W^2) = 2F(X^\Lambda, W^2). \tag{3.186} \]

Near the black hole horizon, the attractor value of $W^2$ obeys $C^2W^2 = 256$, and therefore the exact attractor equations read

\[ p^\Lambda = \text{Re}[CX^\Lambda], \quad q_\Lambda = \text{Re} \left[ CF_\Lambda \left( X^\Lambda, \frac{256}{C^2} \right) \right]. \tag{3.187} \]

This is essentially the only possibility consistent with *symplectic invariance*. It has been then argued that the entropy as a function of the charges is

\[ S_{BH} = \frac{\pi i}{2}(q_\Lambda C \bar{X}^\Lambda - p^\Lambda C \bar{F}_\Lambda) + \frac{\pi}{2} \text{Im}[C^3 \partial_C F], \tag{3.188} \]

where $F_\Lambda, X^\Lambda$ and $C$ are expressed in terms of the charges using (3.187).
Topological Strings

Partition Functions for Black Hole and Topological Strings. The notion of topological string was introduced in [Witten (1990)]. Subsequently a connection between them and superstring was discovered: It was shown in [Bershadsky et al. (1994); Antoniadis et al. (1994)], that the superstring loop corrected $F$–terms (3.185) can be computed as topological string amplitudes. The purpose of this subsection is to translate the super–gravity notation of the previous section to the topological string notation.

The second quantized partition function for the topological string may be written

$$Z_{\text{top}}(t^A, g_{\text{top}}) = \exp \left[ F_{\text{top}}(t^A, g_{\text{top}}) \right],$$

where

$$F_{\text{top}}(t^A, g_{\text{top}}) = \sum_h g_{\text{top}}^{2h-2} F_{\text{top},h}(t^A),$$

and $F_{\text{top},h}$ is the $h$–loop topological string amplitude. The Kähler moduli are expressed in the flat coordinates

$$t^A = \frac{X^A}{X^0} = \theta^A + i r^A,$$

where $r^A$ are the Kähler classes of the Calabi–Yau $M$ and $\theta^A$ are periodic $\theta^A \sim \theta^A + 1$.

We would like to determine relations between super–gravity quantities and topological string quantities. Using the homogeneity property (3.186) and the expansion (3.185), the holomorphic prepotential in super–gravity can be expressed as

$$F(CX^A, 256) = (CX^0)^2 F \left( \frac{X^A}{X^0}, \frac{256}{(CX^0)^2} \right)$$

$$= \sum_{h=0}^{\infty} (CX^0)^{2-2h} f_h(t^A),$$

(3.189)

where $f_h(t^A)$ is related to $F_h(X^A)$ in (3.185) as

$$f_h(t^A) = 16^{2h} F_h \left( \frac{X^A}{X^0} \right).$$

This suggests an identification of the form $f_h(t^A) \sim F_{\text{top},h}(t^A)$ and $g_{\text{top}} \sim (CX^0)^{-1}$. For later purposes, we need precise relations between super–gravity and topological string quantities, including numerical coefficients. These can be determined by studying the limit of a large Calabi–Yau space.
In the super–gravity notation, the genus 0 and 1 terms in the large volume are given by

\[ F(CX^A, 256) = C^2 D_{ABC} \frac{X^A X^B X^C}{X^0} - \frac{1}{6} c_{2A} \frac{X^A}{X^0} + \cdots \]

\[ = (CX^0)^2 D_{ABC} t^A t^B t^C - \frac{1}{6} c_{2A} t^A + \cdots, \]

where \( c_{2A} = \int_M c_2 \wedge \alpha_A \).

with \( c_2 \) being the second Chern class of \( M \), and \( C_{ABC} = -6D_{ABC} \) are the 4-cycle intersection numbers. These terms are normalized so that the mixed entropy \( S_{BH} \) is given by (3.188). On the other hand, the topological string amplitude in this limit is given by

\[ F_{\text{top}} = -\frac{(2\pi)^3 i g_{\text{top}}}{g_{\text{top}}} D_{ABC} t^A t^B t^C - \frac{\pi i}{12} c_{2A} t^A + \cdots \]  

(3.190)

The normalization here is fixed by the holomorphic anomaly equations in Bershadsky et al. (1994), which are nonlinear equations for \( F_{\text{top}} \).

Comparing the one–loop terms in (3.189) and (3.190), which are independent of \( g_{\text{top}} \), we find

\[ F(CX^A, 256) = -\frac{2i}{\pi} F_{\text{top}}(t^A, g_{\text{top}}). \]

Given this, we can compare the genus 0 terms to find

\[ g_{\text{top}} = \pm \frac{4\pi i}{CX^0}. \]

This implies

\[ \ln Z_{BH} = -\pi \ \text{Im} \left[ F(CX^A, 256) \right] = F_{\text{top}} + F_{\text{top}} \quad \text{and} \]

\[ Z_{BH}(\phi^A, \rho^A) = |Z_{\text{top}}(t^A, g_{\text{top}})|^2, \quad \text{with} \]

\[ t^A = \frac{p^A + i\phi^A/\pi}{p^0 + i\phi^0/\pi}, \quad g_{\text{top}} = \pm \frac{4\pi i}{p^0 + i\phi^0/\pi}. \]

**Supergravity Approach to \( Z_{BH} \).** The above relation

\[ Z_{BH} = |Z_{\text{top}}|^2 \]  

(3.191)

can have a simpler super–gravity derivation Ooguri et al. (2004).
A main ingredient in this derivation is the observation that the $N = 2$ super–gravity coupled to vector multiplets can be written as the action

$$S = \int d^4xd^4\theta \text{ (super–volume form)} + \text{h.c.} = \int d^4x\sqrt{-g}R + \ldots, \quad (3.192)$$

where the super–volume form in the above depends non–trivially on curvature of the fields. This reproduces the ordinary action after integrating over $d^4\theta$ and picking up the $\theta^4$ term in the super–volume. In the context of black holes the boundary terms accompanying (3.192) give the classical black hole entropy.

We now become the derivation of (3.191). As was observed in [Bershadsky et al. (1994); Antoniadis et al. (1994)], topological string computes the terms

$$F = \sum_{h=0}^{\infty} \int d^4xd^4\theta F_h(X)(W^2)^g + \text{c.c.} \quad (3.193)$$

There are various terms one can get from the above action after integrating over $d^4\theta$. Let us concentrate on one of the terms which turns out to be the relevant one for us: Take the top components of $X^\Lambda$ and $W^2$, and absorb the $d^4\theta$ integral from the super–volume measure as in (3.192). We will work in the gauge $X^0 \sim 1$ and thus $C \sim 1/g_{\text{top}}^2$. As noted before in the near–horizon black hole geometry in this gauge the top component $W^2 \sim 1/C^2 \sim g_{\text{top}}^2$ and the $X^\Lambda$ are fixed by the attractor mechanism. Thus, we have the black hole free energy

$$\ln Z_{BH} = \sum_{h=0}^{\infty} g_{\text{top}}^{2h} F_{\text{top},h}(X^\Lambda/X^0) \int d^4xd^4\theta + \text{c.c.}$$

$$= \sum_{g=0}^{\infty} (g_{\text{top}}^{2g})^{2h-2} F_{\text{top},h}(X^\Lambda/X^0) + \text{c.c.}$$

$$= 2 \Re F_{\text{top}}, \quad \text{(using } \int d^4xd^4\theta \sim 1/g_{\text{top}}^2).$$

Upon exponentiation this leads to (3.191).

Here we have shown that if we consider one absorption of $\theta^4$ term in (3.193) upon $d^4\theta$ integral we get the desired result. That there be no other terms is not obvious. For example another way to absorb the $\theta$'s would have given the familiar term $R^2 F^{2g-2}$ where $F$ is the graviphoton field. However, such terms do not contribute in the black hole background. It would be
nice to find a simple way to argue why these terms do not contribute and that we are left with this simple absorption of the $\theta$ integrals.

3.4 Non–Quantum Applications of Path Integrals

3.4.1 Stochastic Optimal Control

A path–integral based optimal control model for nonlinear stochastic systems has recently been developed in [Kappen (2005)]. The author addressed the role of noise and the issue of efficient computation in stochastic optimal control problems. He considered a class of nonlinear control problems that can be formulated as a path integral and where the noise plays the role of temperature. The path integral displays symmetry breaking and there exist a critical noise value that separates regimes where optimal control yields qualitatively different solutions. The path integral can be computed efficiently by Monte Carlo integration or by Laplace approximation, and can therefore be used to solve high dimensional stochastic control problems.

Recall that optimal control of nonlinear systems in the presence of noise is a very general problem that occurs in many areas of science and engineering. It underlies autonomous system behavior, such as the control of movement and planning of actions of animals and robots, but also optimization of financial investment policies and control of chemical plants. The problem is stated as: given that the system is in this configuration at this time, what is the optimal course of action to reach a goal state at some future time. The cost of each time course of actions consists typically of a path contribution, that specifies the amount of work or other cost of the trajectory, and an end cost, that specifies to what extend the trajectory reaches the goal state.

Also recall that in the absence of noise, the optimal control problem can be solved in two ways: using (i) the Pontryagin Maximum Principle (PMP, see previous subsection), which represents a pair of ordinary differential equations that are similar to the Hamiltonian equations; or (ii) the Hamilton–Jacobi–Bellman (HJB) equation, which is a partial differential equation (PDE) [Bellman and Kalaba (1964)].

In the presence of Wiener noise, the PMP formalism is replaced by a set of stochastic differential equations (SDEs), which become difficult to solve (compare with [Yong and Zhou (1999)]). The inclusion of noise in the HJB framework is mathematically quite straightforward, yielding the so–called stochastic HJB equation [Stengel (1993)]. However, its solution requires a
discretization of space and time and the computation becomes intractable in both memory requirement and CPU time in high dimensions. As a result, deterministic control can be computed efficiently using the PMP approach, but stochastic control is intractable due to the curse of dimensionality.

For small noise, one expects that optimal stochastic control resembles optimal deterministic control, but for larger noise, the optimal stochastic control can be entirely different from the deterministic control [Russell and Norvig (2003)]. However, there is currently no good understanding how noise affects optimal control.

In this subsection, we address both the issue of efficient computation and the role of noise in stochastic optimal control. We consider a class of nonlinear stochastic control problems, that can be formulated as a statistical mechanics problem. This class of control problems includes arbitrary dynamical systems, but with a limited control mechanism. It contains linear–quadratic [Stengel (1993)] control as a special case. We show that under certain conditions on the noise, the HJB equation can be written as a linear PDE

\[- \partial_t \psi = H \psi, \tag{3.194}\]

with $H$ a (non–Hermitian) operator. Equation (3.194) must be solved subject to a boundary condition at the end time. As a result of the linearity of (3.194), the solution can be obtained in terms of a diffusion process evolving forward in time, and can be written as a path integral. The path–integral has a direct interpretation as a free energy, where noise plays the role of temperature.

This link between stochastic optimal control and a free energy has an immediate consequence that phenomena that allow for a free energy description, typically display phase transitions. [Kappen (2005)] has argued that for stochastic optimal control one can identify a critical noise value that separates regimes where the optimal control has been qualitatively different. He showed how the Laplace approximation can be combined with Monte Carlo sampling to efficiently calculate the optimal control.

3.4.1.1 Path–Integral Formalism

Let $x^i$ be an $nD$ stochastic variable that is subject to the SDE

\[ dx^i = (b^i(x^j, t) + u^i)dt + d\xi^i \tag{3.195} \]
with \(d\xi^i\) being an \(nD\) Wiener process with \(\langle d\xi^i d\xi^j \rangle = \nu_{ij} dt\), and functions \(\nu_{ij}\) independent of \(x^i, u^i\) and time \(t\). The term \(b^i(x^i, t)\) is an arbitrary \(nD\) function of \(x^i\) and \(t\), and \(u^i\) represents an \(nD\) vector of control variables. Given the value of \(x^i\) at an initial time \(t\), the stochastic optimal control problem is to find the control path \(u^i(\cdot)\) that minimizes

\[
C(x^i, t, u^i(\cdot)) = \left\langle \phi(x^i(t_f)) + \int_t^{t_f} d\tau \left( \frac{1}{2} u_i(\tau) R u^i(\tau) + V(x^i(\tau), \tau) \right) \right\rangle_{x^i},
\]

(3.196)

with \(R\) a matrix, \(V(x^i, t)\) a time-dependent potential, and \(\phi(x^i)\) the end cost. The brackets \(\langle\cdot\rangle\) denote expectation value with respect to the stochastic trajectories (3.195) that start at \(x^i\).

One defines the optimal cost-to-go function from any time \(t\) and state \(x^i\) as

\[
J(x^i, t) = \min_{u^i(\cdot)} C(x^i, t, u^i(\cdot)).
\]

(3.197)

\(J\) satisfies the following stochastic HJB equation [Kappen (2005)]

\[
-\partial_t J(x^i, t) = \min_{u^i} \left[ \frac{1}{2} u_i R u^i + V + (b_i + u_i) \partial_{x^i} J(x^i, t) + \frac{1}{2} \nu_{ij} \partial_{x^i} x^j J(x^i, t) \right]
\]

\[
= -\frac{1}{2} R^{-1} \partial_{x^i} J(x^i, t) \partial_{x^i} J + V + b_i \partial_{x^i} J(x^i, t) + \frac{1}{2} \nu_{ij} \partial_{x^i} x^j J(x^i, t),
\]

(3.197)

where \(b_i = (b^i)^T\), and \(u_i = (u^i)^T\), and

\[
u_{ij} \partial_{x^i} x^j J(x^i, t)
\]

is the optimal control at the point \((x^i, t)\). The HJB equation is nonlinear in \(J\) and must be solved with end boundary condition \(J(x^i, t_f) = \phi(x^i)\).

Let us define \(\psi(x^i, t)\) through the Log Transform

\[
J(x^i, t) = -\lambda \log \psi(x^i, t),
\]

(3.199)

and assume that there exists a scalar \(\lambda\) such that

\[
\lambda \delta_{ij} = (R \nu)_{ij},
\]

(3.200)

with \(\delta_{ij}\) the Kronecker delta. In the one dimensional case, such a \(\lambda\) can always be found. In the higher dimensional case, this restricts the matrices \(R \propto (\nu_{ij})^{-1}\). Equation (3.200) reduces the dependence of optimal control on the \(nD\) noise matrix to a scalar value \(\lambda\) that will play the role of
temperature, while \( (3.197) \) reduces to the linear equation \( (3.194) \) with

\[
H = -V + b_i \partial_x^i + \frac{1}{2} \nu_{ij} \partial_x^i \partial_x^j J(x^i, t).
\]

Let \( \rho(y^i, \tau | x^i, t) \) with \( \rho(y^i, t | x^i, t) = \delta(y^i - x^i) \) describe a diffusion process for \( \tau > t \) defined by the Fokker–Planck equation

\[
\partial_\tau \rho = H^\dagger \rho = -V \rho - \partial_x^i (b_i \rho) + \frac{1}{2} \nu_{ij} \partial_x^i x^j J(x^i, t) \rho
\]

with \( H^\dagger \) the Hermitian–conjugate of \( H \). Then

\[
A(\tau) = \int dy^i \rho(y^i, \tau | x^i, t)
\]

is independent of \( \tau \) and in particular \( A(t) = A(t_f) \). It immediately follows that

\[
\psi(x^i, t) = \int dy^i \rho(y^i, t_f | x^i, t) \exp(-\phi(y^i)/\lambda)
\]

We arrive at the important conclusion that \( \psi(x^i, t) \) can be computed either by backward integration using \( (3.194) \) or by forward integration of a diffusion process given by \( (7.19) \).

We can write the integral in \( (3.202) \) as a path integral. Following [Kappen (2005)] we can divide the time interval \( t \rightarrow t_f \) in \( n \) intervals and write

\[
\rho(y^i, t_f | x^i, t) = \prod_{i=1}^n \rho(x^i_i, t_i | x^i_{i-1}, t_{i-1})
\]

and let \( n \rightarrow \infty \). The result is

\[
\psi(x^i, t) = \int [dx^i]_{x^i} \exp \left( -\frac{1}{\lambda} S(x^i(t \rightarrow t_f)) \right)
\]

with \( \int [dx^i]_{x^i} \) an integral over all paths \( x^i(t \rightarrow t_f) \) that start at \( x^i \) and with

\[
S(x^i(t \rightarrow t_f)) = \phi(x^i(t_f)) + \int_t^{t_f} d\tau \left( \frac{1}{2}(\dot{x}^i-b_i(x^i, \tau))R(\dot{x}^i-b'(x^i, \tau)) + V(x^i, \tau) \right)
\]

(3.204) the Action associated with a path. From \( (3.199) \) and \( (3.203) \), the cost–to–go \( J(x, t) \) becomes a log partition sum (i.e., a free energy) with temperature \( \lambda \).

3.4.1.2 Monte Carlo Sampling

The path integral \( (3.203) \) can be estimated by stochastic integration from \( t \) to \( t_f \) of the diffusion process \( (7.19) \) in which particles get annihilated at a rate \( V(x^i, t)/\lambda \) [Kappen (2005)]:

\[
\begin{align*}
\dot{x}^i &= x^i + b^i(x^i, t) dt + d\xi^i, & \text{with probability} & \quad 1 - V dt/\lambda \\
\dot{x}^i &= \uparrow, & \text{with probability} & \quad V dt/\lambda
\end{align*}
\]

(3.205)
where † denotes that the particle is taken out of the simulation. Denote the trajectories by \( x^\alpha_i(t \rightarrow t_f), (\alpha = 1, \ldots, N) \). Then, \( \psi(x^i, t) \) and \( u^i \) are estimated as

\[
\hat{\psi}(x^i, t) = \sum_{\alpha \in \text{alive}} w_\alpha, \quad u^i dt = \frac{1}{\hat{\psi}(x^i, t)} \sum_{\alpha \in \text{alive}} w_\alpha d\xi^i_\alpha(t), \quad (3.206)
\]

with \( w_\alpha = \frac{1}{N} \exp(-\phi(x^i_\alpha(t_f))/\lambda) \),

where 'alive' denotes the subset of trajectories that do not get killed along the way by the † operation. The normalization \( 1/N \) ensures that the annihilation process is properly taken into account. Equation (3.206) states that optimal control at time \( t \) is obtained by averaging the initial directions of the noise component of the trajectories \( d\xi^i_\alpha(t) \), weighted by their success at \( t_f \).

The above sampling procedure can be quite inefficient, when many trajectories get annihilated. One of the simplest procedures to improve it is by importance sampling. We replace the diffusion process that yields \( \rho(y^i, t_f|x^i, t) \) by another diffusion process, that will yield \( \rho'(y^i, t_f|x^i, t) = \exp(-S'/\lambda) \). Then [3.203] becomes,

\[
\psi(x^i, t) = \int [dx^i]_x \exp(-S'/\lambda) \exp(-(S - S')/\lambda).
\]

The idea is to chose \( \rho' \) such as to make the sampling of the path integral as efficient as possible. Following [Kappen (2005)], here we use the Laplace approximation, which is given by the \( k \) deterministic trajectories \( x_\beta(t \rightarrow t_f) \) that minimize the Action

\[
J(x^i, t) \approx -\lambda \log \sum_{\beta=1}^k \exp(-S(x^i_\beta(t \rightarrow t_f))/\lambda).
\]

The Laplace approximation ignores all fluctuations around the modes and becomes exact in the limit \( \lambda \rightarrow 0 \). The Laplace approximation can be computed efficiently, requiring \( \mathcal{O}(n^2m^2) \) operations, where \( m \) is the number of time discretization.

For each Laplace trajectory, we can define a diffusion processes \( \rho'_\beta \) according to [3.205] with \( b'(x^i, t) = x'_\beta(t) \). The estimators for \( \psi \) and \( u^i \) are given again by (3.206), but with weights

\[
w_\alpha = \frac{1}{N} \exp\left(-\left(S(x^i_\alpha(t \rightarrow t_f)) - S'_\beta(x^i_\alpha(t \rightarrow t_f))\right)/\lambda\right).
\]
$S$ is the original Action (3.204) and $S'_\beta$ is the new Action for the Laplace guided diffusion. When there are multiple Laplace trajectories one should include all of these in the sample.

### 3.4.2 Nonlinear Dynamics of Option Pricing

Classical theory of option pricing is based on the results found in 1973 by Black and Scholes \cite{Banos2004} and, independently, Merton \cite{Merton1973}. Their pioneering work starts from the basic assumption that the asset prices follow the dynamics of a particular stochastic process (geometrical Brownian motion), so that they have a lognormal distribution \cite{Hull2000, Paul1999}. In the case of an efficient market with no arbitrage possibilities, no dividends and constant volatilities, they found that the price of each financial derivative is ruled by an ordinary partial differential equation, known as the (Nobel–Prize winning) Black–Scholes–Merton (BSM) formula. In the most simple case of a so-called European option, the BSM equation can be explicitly solved to get an analytical formula for the price of the option \cite{Hull2000, Paul1999}. When we consider other financial derivatives, which are commonly traded in real markets and allow anticipated exercise and/or depend on the history of the underlying asset, the BSM formula fails to give an analytical result. Appropriate numerical procedures have been developed in the literature to price exotic financial derivatives with path–dependent features, as discussed in detail in \cite{Hull2000, Wilmott1993, Potters2001}. The aim of this work is to give a contribution to the problem of efficient option pricing in financial analysis, showing how it is possible to use path integral methods to develop a fast and precise algorithm for the evaluation of option prices.

Following recent studies on the application of the path integral approach to the financial market as appeared in the econophysics literature (see \cite{Matacz2002} for a comprehensive list of references), in \cite{Montagna2002} the authors proposed an original, efficient path integral algorithm to price financial derivatives, including those with path–dependent and early exercise features, and to compare the results with those get with the standard procedures known in the literature.
3.4.2.1 Theory and Simulations of Option Pricing

Classical Theory and Path–Dependent Options

The basic ingredient for the development of a theory of option pricing is a suitable model for the time evolution of the asset prices. The assumption of the BSM model is that the price $S$ of an asset is driven by a Brownian motion and verifies the stochastic differential equation (SDE) [Hull (2000); Paul and Baschnagel (1999)]

$$dS = \mu S dt + \sigma S dw,$$

(3.207)

which, by means of the Itô lemma, can be cast in the form of an arithmetic Brownian motion for the logarithm of $S$

$$d(\ln S) = Adt + \sigma \epsilon dt,$$

(3.208)

where $\sigma$ is the volatility, $A = (\mu - \sigma^2/2)$, $\mu$ is the drift parameter and $\epsilon$ is the realization of a Wiener process. Due to the properties of a Wiener process, (3.208) may be written as

$$d(\ln S) = Adt + \sigma \epsilon \sqrt{dt},$$

(3.209)

where $\epsilon$ follows from a standardized normal distribution with mean 0 and variance 1. Thus, in terms of the logarithms of the asset prices $z' = \ln S'$, $z = \ln S$, the conditional transition probability $p(z'|z)$ to have at the time $t'$ a price $S'$ under the hypothesis that the price was $S$ at the time $t < t'$ is given by [Paul and Baschnagel (1999); Bennati et al. (1999)]

$$p(z'|z) = \frac{1}{\sqrt{2\pi \sigma^2 t'}} \exp \left\{ -\frac{(z' - (z + A(t' - t)))^2}{2\sigma^2(t' - t)} \right\},$$

(3.210)

which is a gaussian distribution with mean $z + A(t' - t)$ and variance $\sigma^2(t' - t)$. If we require the options to be exercised only at specific times $t_i$, $i = 1, \cdots, n$, the asset price, between two consequent times $t_{i-1}$ and $t_i$, will follow (3.209) and the related transition probability will be

$$p(z_i|z_{i-1}) = \frac{1}{\sqrt{2\pi \Delta t \sigma^2}} \exp \left\{ -\frac{(z_i - (z_{i-1} + A\Delta t))^2}{2\sigma^2 \Delta t} \right\},$$

(3.211)

with $\Delta t = t_i - t_{i-1}$. 
A time–evolution model for the asset price is strictly necessary in a theory of option pricing because the fair price at time \( t = 0 \) of an option \( O \), without possibility of anticipated exercise before the expiration date or maturity \( T \) (a so–called **European option**), is given by the scaled expectation value [Hull (2000)]

\[
O(0) = e^{-rT} E[O(T)],
\]

where \( r \) is the risk–free interest and \( E[\cdot] \) indicates the mean value, which can be computed only if a model for the asset underlying the option is understood. For example, the value \( O \) of an European call option at the maturity \( T \) will be \( \max\{ST - X, 0\} \), where \( X \) is the strike price, while for an European put option the value \( O \) at the maturity will be \( \max\{X - ST, 0\} \).

It is worth emphasizing, for what follows, that the case of an European option is particularly simple, since in such a situation the price of the option can be evaluated by means of analytical formulae, which are get by solving the BSM partial differential equation with the appropriate boundary conditions [Hull (2000); Paul and Baschnagel (1999)]. On the other hand, many further kinds of options are present in the financial markets, such as American options (options which can be exercised at any time up to the expiration date) and exotic options [Hull (2000)], i.e., derivatives with complicated payoffs or whose value depend on the whole time evolution of the underlying asset and not just on its value at the end. For such options with path-dependent and early exercise features no exact solutions are available and pricing them correctly is a great challenge.

In the case of options with possibility of anticipated exercise before the expiration date, the above discussion needs to be generalized, by introducing a slicing of the time interval \( T \). Let us consider, for definiteness, the case of an option which can be exercised within the maturity but only at the times \( t_1 = \Delta t, t_2 = 2\Delta t, \ldots, t_n = n\Delta t = T \). At each time slice \( t_{i-1} \), the value \( O_{i-1} \) of the option will be the maximum between its expectation value at the time \( t_i \) scaled with \( e^{-r\Delta t} \) and its value in the case of anticipated exercise \( O^Y_{i-1} \). If \( S_{i-1} \) denotes the price of the underlying asset at the time \( t_{i-1} \), we can thus write for each \( i = 1, \ldots, n \)

\[
O_{i-1}(S_{i-1}) = \max \{ O^Y_{i-1}(S_{i-1}), e^{-r\Delta t} E[O_i|S_{i-1}] \},
\]

where \( E[O_i|S_{i-1}] \) is the conditional expectation value of \( O_i \), i.e., its expectation value under the hypothesis of having the price \( S_{i-1} \) at the time \( t_{i-1} \). In this way, to get the actual price \( O_0 \), it is necessary to proceed back-
ward in time and calculate $O_{n-1}, \ldots, O_1$, where the value $O_n$ of the option at maturity is nothing but $O_Y^n(S_n)$. It is therefore clear that evaluating the price of an option with early exercise features means to simulate the evolution of the underlying asset price (to get the $O_Y^n$) and to calculate a (usually large) number of expectation conditional probabilities.

**Standard Numerical Procedures**

To value derivatives when analytical formulae are not available, appropriate numerical techniques have to be advocated. They involve the use of Monte Carlo (MC) simulation, binomial trees (and their improvements) and finite–difference methods [Hull (2000); Wilmott et al. (1993)].

A natural way to simulate price paths is to discretize (3.209) as

$$\ln S(t + \Delta t) - \ln S(t) = A \Delta t + \sigma \epsilon \sqrt{\Delta t},$$

or, equivalently,

$$S(t + \Delta t) = S(t) \exp [A \Delta t + \sigma \epsilon \sqrt{\Delta t}], \quad (3.214)$$

which is correct for any $\Delta t > 0$, even if finite. Given the spot price $S_0$, i.e., the price of the asset at time $t = 0$, one can extract from a standardized normal distribution a value $\epsilon_k, (k = 1, \ldots, n)$ for the random variable $\epsilon$ to simulate one possible path followed by the price by means of (3.214):

$$S(k \Delta t) = S((k-1) \Delta t) \exp [A \Delta t + \sigma \epsilon_k \sqrt{\Delta t}].$$

Iterating the procedure $m$ times, one can simulate $m$ price paths $\{(S_0, S_1^{(j)}, S_2^{(j)}, \ldots, S_T^{(j)} \equiv S_T) : j = 1, \ldots, m\}$ and evaluate the price of the option. In such a MC simulation of the stochastic dynamics of asset price (Monte Carlo random walk) the mean values $E[O_i|S_{i-1}], i = 1, \ldots, n$ are given by

$$E[O_i|S_{i-1}] = \frac{O_i^{(1)} + O_i^{(2)} + \cdots + O_i^{(m)}}{m},$$

with no need to calculate transition probabilities because, through the extraction of the possible $\epsilon$ values, the paths are automatically weighted according to the probability distribution function of (3.211). Unfortunately, this method leads to an estimated value whose numerical error is proportional to $m^{-1/2}$. Thus, even if it is powerful because of the possibility to control the paths and to impose additional constrains (as it is usually
required by exotic and path-dependent options), the MC random walk is extremely time consuming when precise predictions are required and appropriate variance reduction procedures have to be used to save CPU time [Hull (2000)]. This difficulty can be overcome by means of the method of the binomial trees and its extensions (see [Hull (2000) and references therein]), whose main idea stands in a deterministic choice of the possible paths to limit the number of intermediate points. At each time step the price \( S_i \) is assumed to have only two choices: increase to the value \( uS_i, u > 1 \) or decrease to \( dS_i, 0 < d < 1 \), where the parameters \( u \) and \( d \) are given in terms of \( \sigma \) and \( \Delta t \) in such a way to give the correct values for the mean and variance of stock price changes over the time interval \( \Delta t \). Also finite difference methods are known in the literature [Hull (2000)] as an alternative to time-consuming MC simulations. They give the value of the derivative by solving the differential equation satisfied by the derivative, by converting it into a difference equation. Although tree approaches and finite difference methods are known to be faster than the MC random walk, they are difficult to apply when a detailed control of the history of the derivative is required and are also computationally time consuming when a number of stochastic variables is involved [Hull (2000)]. It follows that the development of efficient and fast computational algorithms to price financial derivatives is still a key issue in financial analysis.

3.4.2.2 Option Pricing via Path Integrals

Recall that the path integral method is an integral formulation of the dynamics of a stochastic process. It is a suitable framework for the calculation of the transition probabilities associated to a given stochastic process, which is seen as the convolution of an infinite sequence of infinitesimal short-time steps [Bennati et al. (1999)]. For the problem of option pricing, the path–integral method can be employed for the explicit calculation of the expectation values of the quantities of financial interest, given by integrals of the form [Bennati et al. (1999)]

\[
E[O_i|S_{i-1}] = \int dz_i p(z_i|z_{i-1}) O_i(e^{z_i}),
\]

(3.215)

where \( z = \ln S \) and \( p(z_i|z_{i-1}) \) is the transition probability. \( E[O_i|S_{i-1}] \) is the conditional expectation value of some functional \( O_i \) of the stochastic process. For example, for an European call option at the maturity \( T \) the quantity of interest will be \( \max\{S_T - X, 0\} \), \( X \) being the strike
price. As already emphasized, and discussed in the literature [Hull (2000); Wilmott et al. (1993); Potters et al. (2001); Rosa-Clot and Taddei (2002); Matacz (2002)], the computational complexity associated to this calculation is generally great: in the case of exotic options, with path-dependent and early exercise features, integrals of the type (3.215) cannot be analytically solved. As a consequence, we demand two things from a path integral framework: a very quick way to estimate the transition probability associated to a stochastic process (3.209) and a clever choice of the integration points with which evaluate the integrals (3.215). In particular, our aim is to develop an efficient calculation of the probability distribution without losing information on the path followed by the asset price during its time evolution.

Transition Probability

The probability distribution function related to a SDE verifies the Chapman–Kolmogorov equation [Paul and Baschnagel (1999)]

\[ p(z''|z') = \int dz p(z''|z)p(z|z'), \]  

(3.216)

which states that the probability (density) of a transition from the value \( z' \) (at time \( t' \)) to the value \( z'' \) (at time \( t'' \)) is the ‘summation’ over all the possible intermediate values \( z \) of the probability of separate and consequent transitions \( z' \rightarrow z, z \rightarrow z'' \). As a consequence, if we consider a finite time interval \( [t', t''] \) and we apply a time slicing, by considering \( n+1 \) subintervals of length \( \Delta t = (t'' - t')/n + 1 \), we can write, by iteration of (3.216)

\[ p(z''|z') = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} dz_1 \cdots dz_n p(z''|z_n)p(z_n|z_{n-1}) \cdots p(z_1|z'), \]

(3.217)

which, thanks to (3.210), can be written as [Montagna et al. (2002)]

\[ \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} dz_1 \cdots dz_n \frac{1}{(2\pi\sigma^2\Delta t)^{n+1}} \exp \left\{ -\frac{1}{2\sigma^2\Delta t} \sum_{k=1}^{n+1} [z_k - (z_{k-1} + A\Delta t)]^2 \right\}. \]

In the limit \( n \rightarrow \infty, \Delta t \rightarrow 0 \) such that \((n + 1)\Delta t = (t'' - t')\) (infinite sequence of infinitesimal time steps), the expression (3.217), as explicitly shown in [Bennati et al. (1999)], exhibits a Lagrangian structure and it is
possible to express the transition probability in the path integral formalism as a convolution of the form \[Bennati \text{ et al. (1999)}\]

\[
p(z'', t''| z', t') = \int_{C} D[\sigma^{-1} \ddot{z}] \exp \left\{ - \int_{t'}^{t''} L(\dot{\ddot{z}}, \ddot{z}; \tau) d\tau \right\},
\]

where \( L \) is the Lagrangian, given by

\[
L(\dot{\ddot{z}}(\tau), \ddot{z}(\tau); \tau) = \frac{1}{2} \sigma^2 \left[ \dot{\ddot{z}}(\tau) - A \right]^2,
\]

and the integral is performed (with functional measure \( D[\cdot] \)) over the paths \( \ddot{z} \) belonging to \( C \), i.e., all the continuous functions with constrains \( \ddot{z}(t') \equiv z', \ddot{z}(t'') \equiv z'' \). As carefully discussed in \[Bennati \text{ et al. (1999)}\], a path integral is well defined only if both a continuous formal expression and a discretization rule are given. As done in many applications, the Itô prescription is adopted here (see subsection 3.1.2.5 above).

A first, na"{i}ve evaluation of the transition probability (3.217) can be performed via Monte Carlo simulation, by writing (3.217) as

\[
p(z'', t''| z', t') = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \prod_{i=1}^{n} dg_i \frac{1}{\sqrt{2\pi \sigma^2 \Delta t}} \exp \left\{ - \frac{1}{2\sigma^2 \Delta t} [z'' - (z_{n} + A\Delta t)]^2 \right\},
\]

in terms of the variables \( g_i \) defined by the relation

\[
dg_k = \frac{dz_k}{\sqrt{2\pi \sigma^2 \Delta t}} \exp \left\{ - \frac{1}{2\sigma^2 \Delta t} [z_k - (z_{k-1} + A\Delta t)]^2 \right\},
\]

and extracting each \( g_i \) from a gaussian distribution of mean \( z_{k-1} + A\Delta t \) and variance \( \sigma^2 \Delta t \). However, as we will see, this method requires a large number of calls to get a good precision. This is due to the fact that each \( g_i \) is related to the previous \( g_{i-1} \), so that this implementation of the path integral approach can be seen to be equivalent to a na"{i}ve MC simulation of random walks, with no variance reduction.

By means of appropriate manipulations \[Schulman (1981)\] of the integrand entering (3.217), it is possible, as shown in the following, to get a path integral expression which will contain a factorized integral with a constant kernel and a consequent variance reduction. If we define \( z'' = z_{n+1} \) and \( y_k = z_k - k A \Delta t, k = 1, \ldots, n \), we can express the transition probability

\[
\int_{C} D[\sigma^{-1} \ddot{z}] \exp \left\{ - \int_{t'}^{t''} L(\dot{\ddot{z}}, \ddot{z}; \tau) d\tau \right\},
\]
distribution as
\[ \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} dy_1 \cdots dy_n \frac{1}{\sqrt{(2\pi \sigma^2 \Delta t)^{n+1}}} \exp \left\{ -\frac{1}{2 \sigma^2 \Delta t} \sum_{k=1}^{n+1} (y_k - y_{k-1})^2 \right\}, \] (3.220)
in order to get rid of the contribution of the drift parameter. Now let us extract from the argument of the exponential function a quadratic form
\[ \sum_{k=1}^{n+1} (y_k - y_{k-1})^2 = y_0^2 - 2y_1y_0 + y_1^2 + y_1^2 - 2y_1y_2 + \cdots + y_{n+1}^2, \] (3.221)
by introducing the \( n \times n \) matrix \( M \) defined as \cite{Montagna et al. (2002)}
\[ y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad M = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & -1 & 2 & -1 \\ 0 & \cdots & \cdots & -1 & 2 \end{pmatrix} \] (3.222)
where \( M \) is a real, symmetric, non singular and tridiagonal matrix. In terms of the eigenvalues \( m_i \) of the matrix \( M \), the contribution in (3.221) can be written as
\[ y^t M y = \sum_{i=1}^{n} m_i w_i^2, \] (3.223)
by introducing the orthogonal matrix \( O \) which diagonalizes \( M \), with \( w_i = O_{ij} y_j \). Because of the orthogonality of \( O \), the Jacobian
\[ J = \det \left| \frac{dw_i}{dy_k} \right| = \det |O_{ki}|, \]
of the transformation \( y_k \rightarrow w_k \) equals 1, so that \( \prod_{i=1}^{n} dw_i = \prod_{i=1}^{n} dy_i \). After some algebra, (3.221) can be written as
\[ \sum_{k=1}^{n+1} (y_k - y_{k-1})^2 = \sum_{i=1}^{n} m_i w_i^2 + y_0^2 - 2y_1y_0 + y_{n+1}^2 - 2y_ny_{n+1} = \]
\[
\sum_{i=1}^{n} m_i \left[ w_i - \frac{(y_0O_{11} + y_{n+1}O_{ni})}{m_i} \right]^2 + y_0^2 + y_{n+1}^2 - \sum_{i=1}^{n} \left( \frac{(y_0O_{11} + y_{n+1}O_{ni})^2}{m_i} \right).
\]

Now, if we introduce new variables \( h_i \) obeying the relation

\[
dh_i = \sqrt{\frac{m_i}{2\pi\sigma^2\Delta t}} \exp \left\{ -\frac{m_i}{2\pi\sigma^2\Delta t} \left( w_i - \frac{(y_0O_{11} + y_{n+1}O_{ni})}{m_i} \right)^2 \right\} dw_i,
\]

(3.225)

it is possible to express the finite–time probability distribution \( p(z''|z') \) as

[Montagna et al. (2002)]

\[
\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \prod_{i=1}^{n} dy_i \frac{1}{\sqrt{(2\pi\sigma^2\Delta t)^n+1}} \exp \left\{ -\frac{1}{2\pi^2\Delta t} \sum_{k=1}^{n+1} [y_k - y_{k-1}]^2 \right\}
\]

\[
= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \prod_{i=1}^{n} dw_i \frac{1}{\sqrt{(2\pi\sigma^2\Delta t)^n+1}} \exp \left\{ -\frac{1}{2\pi^2\Delta t} \sum_{i=1}^{n} \left[ m_i \left( w_i - \frac{(y_0O_{11} + y_{n+1}O_{ni})}{m_i} \right)^2 - \frac{(y_0O_{11} + y_{n+1}O_{ni})^2}{m_i} \right] \right\}
\]

\[
= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \prod_{i=1}^{n} dh_i \frac{1}{2\pi\sigma^2\Delta t \det(M)} \times \exp \left\{ -\frac{1}{2\pi^2\Delta t} \left[ y_0^2 + y_{n+1}^2 + \sum_{i=1}^{n} \left( \frac{(y_0O_{11} + y_{n+1}O_{ni})^2}{m_i} \right) \right] \right\}.
\]

The probability distribution function, as given by (3.226), is an integral whose kernel is a constant function (with respect to the integration variables) and which can be factorized into the \( n \) integrals

\[
\int_{-\infty}^{+\infty} dh_i \exp \left\{ -\frac{1}{2\pi^2\Delta t} \left( \frac{(y_0O_{11} + y_{n+1}O_{ni})^2}{m_i} \right) \right\},
\]

(3.227)

given in terms of the \( h_i \), which are gaussian variables that can be extracted from a normal distribution with mean \( (y_0O_{11} + y_{n+1}O_{ni})^2/m_i \) and variance \( \sigma^2\Delta t/m_i \). Differently to the first, naïve implementation of the path integral, now each \( h_i \) is no longer dependent on the previous \( h_{i-1} \), and importance sampling over the paths is automatically accounted for.

It is worth noticing that, by means of the extraction of the random variables \( h_i \), we are creating price paths, since at each intermediate time \( t_i \)
the asset price is given by
\[ S_i = \exp \left\{ \sum_{k=1}^{n} O_{ik} h_k + iA\Delta t \right\}. \] (3.228)

Therefore, this path integral algorithm can be easily adapted to the cases in which the derivative to be valued has, in the time interval \([0,T]\), additional constraints, as in the case of interesting path-dependent options, such as Asian and barrier options [Hull (2000)].

Integration Points

The above illustrated method represents a powerful and fast tool to calculate the transition probability in the path integral framework and it can be employed if we need to value a generic option with maturity \(T\) and with possibility of anticipated exercise at times \(t_i = i\Delta t\ (n\Delta t = T)\) [Montagna et al. (2002)]. As a consequence of this time slicing, one must numerically evaluate \(n - 1\) mean values of the type (9), in order to check at any time \(t_i\), and for any value of the stock price, whether early exercise is more convenient with respect to holding the option for a future time. To keep under control the computational complexity and the time of execution, it is mandatory to limit as far as possible the number of points for the integral evaluation. This means that we would like to have a linear growth of the number of integration points with the time. Let us suppose to evaluate each mean value
\[ E[O_i|S_{i-1}] = \int dz_i p(z_i|z_{i-1}) O_i(e^{z_i}), \]
with \(p\) integration points, i.e., considering only \(p\) fixed values for \(z_i\). To this end, we can create a grid of possible prices, according to the dynamics of the stochastic process as given by (3.209)
\[ z(t + \Delta t) - z(t) = \ln S(t + \Delta t) - \ln S(t) = A\Delta t + \epsilon \sigma \sqrt{\Delta t}. \] (3.229)

Starting from \(z_0\), we thus evaluate the expectation value \(E[O_1|S_0]\) with \(p = 2m+1, m \in \mathbb{N}\) values of \(z_1\) centered on the mean value \(E[z_1] = z_0 + A\Delta t\) and which differ from each other of a quantity of the order of \(\sigma \sqrt{\Delta t}\)
\[ z_1^j = z_0 + A\Delta t + j\sigma \sqrt{\Delta t}, \quad (j = -m, \ldots, +m). \]

Going on like this, we can evaluate each expectation value \(E[O_2|z_1^j]\) get from each one of the \(z_1^j\)’s created above with \(p\) values for \(z_2\) centered around the
mean value

\[ E[z_2|z_1^1] = z_1^1 + A\Delta t = z_0 + 2A\Delta t + j\sigma \sqrt{\Delta t}. \]

Iterating the procedure until the maturity, we create a deterministic
grid of points such that, at a given time \( t_i \), there are \( (p - 1)i + 1 \) values of
\( z_i \), in agreement with the request of linear growth. This procedure of se-
lection of integration points, together with the calculation of the transition
probability previously described, is the basis of the path integral simulation
of the price of a generic option.

By applying the results derived above, we have at disposal an efficient
path integral algorithm both for the calculation of transition probabilities
and the evaluation of option prices. In [Montagna et al. (2002)] the ap-
plication of the above path–integral method to European and American
options in the BSM model was illustrated and comparisons with the results
were get with the standard procedures known in the literature were shown.
First, the path integral simulation of the probability distribution of the
logarithm of the stock prices, \( p(\ln S) \), as a function of the logarithm of the
stock price, for a BSM–like stochastic model, was given by (3.208). Once
the transition probability has been computed, the price of an option could
be computed in a path integral approach as the conditional expectation
value of a given functional of the stochastic process. For example, the price
of an European call option was given by

\[ C = e^{-r(T-t)} \int_{-\infty}^{+\infty} dz_f p(z_f, T|z_i, t) \max[e^{z_f} - X, 0], \]  

(3.230)

while for an European put it will be

\[ P = e^{-r(T-t)} \int_{-\infty}^{+\infty} dz_f p(z_f, T|z_i, t) \max[X - e^{z_f}, 0], \]  

(3.231)

where \( r \) is the risk–free interest rate. Therefore just 1D integrals need to be
evaluated and they can be precisely computed with standard quadrature
rules.

3.4.2.3 Continuum Limit and American Options

In the specific case of an American option, the possibility of exercise at any
time up to the expiration date allows to develop, within the path integral
formalism, a specific algorithm, which, as shown in the following, is precise
and very quick [Montagna et al. (2002)].
Given the time slicing considered above, the case of American options
requires the limit $\Delta t \to 0$ which, putting $\sigma \to 0$, leads to a delta–like
transition probability

$$p(z, t + \Delta t | z_t, t) \approx \delta(z - z_t - A \Delta t).$$

This means that, apart from volatility effects, the price $z_i$ at time $t_i$
will have a value remarkably close to the expected value $\bar{z} = z_{i-1} + A \Delta t,$
given by the drift growth. In order to take care of the volatility effects, a possible
solution is to estimate the integral of interest, i.e.,

$$E[O_i | S_{i-1}] = \int_{-\infty}^{+\infty} dz p(z | z_{i-1}) O_i(e^z),$$

(3.232)

by inserting in (3.232) the analytical expression for the $p(z | z_{i-1})$ transition
probability

$$p(z | z_{i-1}) = \frac{1}{\sqrt{2\pi \Delta t \sigma^2}} \exp \left\{ - \frac{(z - z_{i-1} - A \Delta t)^2}{2 \sigma^2 \Delta t} \right\}$$

$$= \frac{1}{\sqrt{2\pi \Delta t \sigma^2}} \exp \left\{ - \frac{(z - \bar{z})^2}{2 \sigma^2 \Delta t} \right\},$$

together with a Taylor expansion of the kernel function $O_i(e^z) = f(z)$
around the expected value $\bar{z}$. Hence, up to the second–order in $z - \bar{z}$, the
kernel function becomes

$$f(z) = f(\bar{z}) + (z - \bar{z}) f'(\bar{z}) + \frac{1}{2} f''(\bar{z})(z - \bar{z})^2 + O((z - \bar{z})^3),$$

which induces

$$E[O_i | S_{i-1}] = f(\bar{z}) + \frac{\sigma^2}{2} f''(\bar{z}) + \ldots,$$

since the first derivative does not give contribution to (3.232), being the
integral of an odd function over the whole $z$ range. The second derivative
can be numerically estimated as

$$f''(\bar{z}) = \frac{1}{\delta_\sigma} [f(\bar{z} + \delta_\sigma) - 2f(\bar{z}) + f(\bar{z} - \delta_\sigma)],$$

with $\delta_\sigma = O(\sigma \sqrt{\Delta t})$, as dictated by the dynamics of the stochastic process.
3.4.3 Dynamics of Complex Networks

Recall that many systems in nature, such as neural nets, food webs, metabolic systems, co-authorship of papers, the worldwide web, etc. can be represented as complex networks, or small-world networks (see, e.g., [Watts and Strogatz (1998); Dorogovtsev and Mendes (2003)]). In particular, it has been recognized that many networks have scale-free topology; the distribution of the degree obeys the power law, $P(k) \sim k^{-\gamma}$. The study of the scale-free network now attracts the interests of many researchers in mathematics, physics, engineering and biology [Ichinomiya (2004)].

Another important aspect of complex networks is their dynamics, describing e.g., the spreading of viruses in the Internet, change of populations in a food web, and synchronization of neurons in a brain. In particular, [Ichinomiya (2004)] studied the synchronization of the random network of oscillators. His work follows the previous studies (see [Strogatz (2000)]) that showed that mean-field type synchronization, that Kuramoto observed in globally-coupled oscillators [Kuramoto (1984)], appeared also in the small-world networks.

3.4.3.1 Continuum Limit of the Kuramoto Net

Ichinomiya started with the standard network with $N$ nodes, described by a variant of the Kuramoto model. Namely, at each node, there exists an oscillator and the phase of each oscillator $\theta_i$ is evolving according to

$$\dot{\theta}_i = \omega_i + K \sum_j a_{ij} \sin(\theta_j - \theta_i), \quad (3.233)$$

where $K$ is the coupling constant, $a_{ij}$ is 1 if the nodes $i$ and $j$ are connected, and 0 otherwise; $\omega_i$ is a random number, whose distribution is given by the function $N(\omega)$.

For the analytic study, it is convenient to use the continuum limit equation. We define $P(k)$ as the distribution of nodes with degree $k$, and $\rho(k, \omega; t, \theta)$ the density of oscillators with phase $\theta$ at time $t$, for given $\omega$ and $k$. We assume that $\rho(k, \omega; t, \theta)$ is normalized as

$$\int_0^{2\pi} \rho(k, \omega; t, \theta) d\theta = 1.$$

For simplicity, we also assume $N(\omega) = N(-\omega)$. Thus, we suppose that the collective oscillation corresponds to the stable solution, $\dot{\rho} = 0$. 
Now we construct the continuum limit equation for the network of oscillators. The evolution of $\rho$ is determined by the \textit{continuity equation} $\partial_t \rho = -\partial_\theta (\rho v)$, where $v$ is defined by the continuum limit of the r.h.s of (3.233).

Because one randomly selected edge connects to the node of degree $k$, frequency $\omega$, phase $\theta$ with the probability $kP(k)N(\omega)\rho(k, \omega; t, \theta) / \int dk kP(k)$, $\rho(k, \omega; t, \theta)$ obeys the equation

$$\partial_t \rho(k, \omega; t, \theta) = -\partial_\theta [\rho(k, \omega; t, \theta) (\omega + K k \int d\omega' \int dk' \int d\theta' N(\omega') P(k') \rho(k', \omega'; t, \theta') \sin(\theta - \theta') \int dk'' P(k'')]].$$

The mean–field solution of this equation was studied by Ichinomiya (2004).

### 3.4.3.2 Path–Integral for Complex Networks

Recently, Ichinomiya (2005) introduced the \textit{path–integral} (see subsection 4.4.6 above) \textit{approach} in studying the dynamics of complex networks. He considered the stochastic generalization of the Kuramoto network (3.233), given by

$$\dot{x}_i = f_i(x_i) + \sum_{j=1}^N a_{ij} g(x_i, x_j) + \xi_i(t),$$

(3.234)

where $f_i = f_i(x_i)$ and $g_{ij} = g(x_i, x_j)$ are functions of network activations $x_i$, $\xi_i(t)$ is a \textit{random force} that satisfies $\langle \xi_i(t) \rangle = 0$, $\langle \xi_i(t) \xi_j(t') \rangle = \delta_{ij} \delta(t - t') \sigma^2$.

He assumed $x_i = x_{i,0}$ at $t = 0$. In order to discuss the dynamics of this system, he introduced the so-called Matrin–Siggia–Rose (MSR) \textit{generating functional} $Z$ given by Dominicis (1978)

$$Z[\{l_{ik}\}, \{\bar{l}_{ik}\}] = \left( \frac{1}{\pi} \right)^{NN} \left\{ \prod_{i=1}^N \prod_{k=0}^{N_t} dx_{ik} d\bar{x}_{ik} e^{-S} \exp(l_{ik} x_{ik} + \bar{l}_{ik} \bar{x}_{ik}) \right\},$$

where the \textit{action} $S$ is given by

$$S = \sum_{ik} \left[ \frac{\sigma^2 \Delta t}{2} \bar{x}_{ik}^2 + i \bar{x}_{ik} \{ x_{ik} - x_{i,k-1} - \Delta t (f_i(x_{i,k-1}) + \sum_j a_{ij} g(x_{i,k-1}, x_{j,k-1})) \} \right].$$
and \( \langle \cdots \rangle \) represents the average over the ensemble of networks. \( J \) is the functional Jacobian term,

\[
J = \exp \left( -\frac{\Delta t}{2} \sum_{ijk} \frac{\partial (f_i(x_{ik}) + a_{ij}g(x_{ik}, x_{jk}))}{\partial x_{ik}} \right).
\]

Ichinomiya considered such a form of the network model in which

\[
a_{ij} = \begin{cases} 
1 & \text{with probability } p_{ij}, \\
0 & \text{with probability } 1 - p_{ij}.
\end{cases}
\]

Note that \( p_{ij} \) can be a function of variables such as \( i \) or \( j \). For example, in the 1D chain model, \( p_{ij} \) is 1 if \( |i - j| = 1 \), else it is 0. The average over all networks can be expressed as

\[
\langle \exp \left[ \sum_{ik} i \Delta t \bar{x}_{ik} \sum_{j} a_{ij}g(x_{i,k-1}, x_{j,k-1}) \right]\rangle = \prod_{ij} p_{ij} \exp \left\{ \sum_{k} i \Delta t \bar{x}_{ik}g(x_{i,k-1}, x_{j,k-1}) \right\} + 1 - p_{ij},
\]

so we get

\[
\langle e^{-S} \rangle = \exp(-S_0) \prod_{ij} p_{ij} \exp \left\{ \sum_{k} i \Delta t \bar{x}_{ik}g(x_{i,k-1}, x_{j,k-1}) \right\} + 1 - p_{ij},
\]

where \( S_0 = \sum_{ik} \frac{\sigma^2 \Delta t}{2} \bar{x}_{ik}^2 + i \bar{x}_{ik} \{ x_{ik} - x_{i,k-1} - \Delta t f_i(x_{i,k-1}) \} \).

This expression can be applied to the dynamics of any complex network model. Ichinomiya [2005] applied this model to analysis of the Kuramoto transition in random sparse networks.

### 3.4.4 Cerebellum as a Neural Path–Integral

Recall that human motion is naturally driven by synergistic action of more than 600 skeletal muscles. While the muscles generate driving torques in the moving joints, subcortical neural system performs both local and global (loco)motion control: first reflexly controlling contractions of individual muscles, and then orchestrating all the muscles into synergetic actions in order to produce efficient movements. While the local reflex control of
individual muscles is performed on the spinal control level, the global integration of all the muscles into coordinated movements is performed within the cerebellum.

All hierarchical subcortical neuro-muscular physiology, from the bottom level of a single muscle fiber, to the top level of cerebellar muscular synergy, acts as a temporal $\langle \text{out} | \text{in} \rangle$ reaction, in such a way that the higher level acts as a command/control space for the lower level, itself representing an abstract image of the lower one:

1. At the muscular level, we have excitation-contraction dynamics [Hatze (1977a); Hatze (1978); Hatze (1977b)], in which $\langle \text{out} | \text{in} \rangle$ is given by the following sequence of nonlinear diffusion processes: neural-action-potential $\rightarrow$ synaptic-potential $\rightarrow$ muscular-action-potential $\rightarrow$ excitation-contraction-coupling $\rightarrow$ muscle-tension-generating [Ivancevic (1991); Ivancevic and Ivancevic (2006a)]. Its purpose is the generation of muscular forces, to be transferred into driving torques within the joint anatomical geometry.

2. At the spinal level, $\langle \text{out} | \text{in} \rangle$ is given by autogenetic-reflex stimulus-response control [Houk (1979)]. Here we have a neural image of all individual muscles. The main purpose of the spinal control level is to give both positive and negative feedbacks to stabilize generated muscular forces within the 'homeostatic' (or, more appropriately, 'homeokinetic') limits. The individual muscular actions are combined into flexor-extensor (or agonist-antagonist) pairs, mutually controlling each other. This is the mechanism of reciprocal innervation of agonists and inhibition of antagonists. It has a purely mechanical purpose to form the so-called equivalent muscular actuators (EMAs), which would generate driving torques $T_i(t)$ for all movable joints.

3. At the cerebellar level, $\langle \text{out} | \text{in} \rangle$ is given by sensory-motor integration [Houk et al. (1996)]. Here we have an abstracted image of all autogenetic reflexes. The main purpose of the cerebellar control level is integration and fine tuning of the action of all active EMAs into a synchronized movement, by supervising the individual autogenetic reflex circuits. At the same time, to be able to perform in new and unknown conditions, the cerebellum is continuously adapting its own neural circuitry by unsupervised (self-organizing) learning. Its action is subconscious and automatic, both in humans and in animals.

Naturally, we can ask the question: Can we assign a single $\langle \text{out} | \text{in} \rangle$
Quantum Leap

measure to all these neuro–muscular stimulus–response reactions? We think that we can do it; so in this Letter, we propose the concept of adaptive sensory–motor transition amplitude as a unique measure for this temporal $<\text{out}|\text{in}>$ relation. Conceptually, this $<\text{out}|\text{in}>$ amplitude can be formulated as the ‘neural path integral’:

$$<\text{out}|\text{in}> \equiv \langle \text{sensory}|\text{motor} \rangle = \langle \text{motor}|\text{sensory} \rangle = \int D[w,x] e^{iS[x]},$$

(3.235)

Here, the integral is taken over all activated (or, ‘fired’) neural pathways $x^i = x^i(t)$ of the cerebellum, connecting its input sensory–state with its output motor–state, symbolically described by adaptive neural measure $D[w,x]$, defined by the weighted product (of discrete time steps)

$$D[w,x] = \lim_{n \to \infty} \prod_{t=1}^n w^i(t) dx^i(t),$$

in which the synaptic weights $w^i = w^i(t)$, included in all active neural pathways $x^i = x^i(t)$, are updated by the unsupervised Hebbian–like learning rule (7.63), namely

$$w^i(t+1) = w^i(t) + \sigma \frac{\eta}{\eta} (w^i_d(t) - w^i_a(t)),$$

(3.236)

where $\sigma = \sigma(t)$, $\eta = \eta(t)$ represent local neural signal and noise amplitudes, respectively, while superscripts $d$ and $a$ denote desired and achieved neural states, respectively. Theoretically, equations (3.235–3.236) define an $\infty$–dimensional neural network. Practically, in a computer simulation we can use $10^7 \leq n \leq 10^8$, roughly corresponding to the number of neurons in the cerebellum.

The exponent term $S[x]$ in equation (3.235) represents the autogenetic–reflex action, describing reflexly–induced motion of all active EMAs, from their initial stimulus–state to their final response–state, along the family of extremal (i.e., Euler–Lagrangian) paths $x^i_{\min}(t)$. ($S[x]$ is properly derived in (3.239–3.240) below.)

3.4.4.1 Spinal Autogenetic Reflex Control

Recall (from Introduction) that at the spinal control level we have the autogenetic reflex motor servo [Houk (1979)], providing the local, reflex feedback loops for individual muscular contractions. A voluntary contraction force $F$ of human skeletal muscle is reflexly excited (positive feedback $+F^{-1}$)
by the responses of its *spindle receptors* to stretch and is reflexly inhibited (negative feedback \(-F^{-1}\)) by the responses of its *Golgi tendon organs* to contraction. Stretch and unloading reflexes are mediated by combined actions of several autogenetic neural pathways, forming the *motor servo*.

In other words, branches of the afferent fibers also synapse with interneurons that inhibit motor neurons controlling the antagonistic muscles – *reciprocal inhibition*. Consequently, the stretch stimulus causes the antagonists to relax so that they cannot resist the shortening of the stretched muscle caused by the main reflex arc. Similarly, firing of the Golgi tendon receptors causes inhibition of the muscle contracting too strong and simultaneous *reciprocal activation* of its antagonist. Both mechanisms of reciprocal inhibition and activation performed by the autogenetic circuits \(+F^{-1}\) and \(-F^{-1}\), serve to generate the well–tuned EMA–driving torques \(T_i\).

Now, once we have properly defined the symplectic musculo–skeletal dynamics [Ivancevic (2004)] on the biomechanical (momentum) phase–space manifold \(T^*M^N\), we can proceed in formalizing its hierarchical subcortical neural control. By introducing the *coupling Hamiltonians* \(H^m = H^m(q,p)\), selectively corresponding only to the \(M \leq N\) active joints, we define the *affine Hamiltonian control function* \(H_{aff} : T^*M^N \rightarrow \mathbb{R}\), in local canonical coordinates on \(T^*M^N\) given by (adapted from [Ivancevic and Ivancevic (2006b)] for the biomechanical purpose)

\[
H_{aff}(q,p) = H_0(q,p) - H^m(q,p)T_m, \quad (m = 1, \ldots, M \leq N),
\]

where \(T_m = T_m(t,q,p)\) are affine feedback torque one–forms, different from the initial driving torques \(T_i\) acting in all the joints. Using the affine Hamiltonian function \(3.237\), we get the *affine Hamiltonian servo–system* [Ivancevic (2004)].

\[
\begin{align*}
q^i &= \frac{\partial H_0(q,p)}{\partial p_i} - \frac{\partial H^m(q,p)}{\partial p_i}T_m, \\
p_i &= -\frac{\partial H_0(q,p)}{\partial q^i} + \frac{\partial H^m(q,p)}{\partial q^i}T_m, \\
q^i(0) &= q_0^i, \quad p_i(0) = p_i^0, \quad (i = 1, \ldots, N; \quad m = 1, \ldots, M \leq N).
\end{align*}
\]

The affine Hamiltonian control system \(3.238\) gives our formal description for the autogenetic spinal motor–servo for all \(M \leq N\) activated (i.e., working) EMAs.
3.4.4.2 Cerebellum – The Comparator

Having, thus, defined the spinal reflex control level, we proceed to model the top subcortical commander/controller, the cerebellum. It is a brain region anatomically located at the bottom rear of the head (the hindbrain), directly above the brainstem, which is important for a number of subconscious and automatic motor functions, including motor learning. It processes information received from the motor cortex, as well as from proprioceptors and visual and equilibrium pathways, and gives ‘instructions’ to the motor cortex and other subcortical motor centers (like the basal nuclei), which result in proper balance and posture, as well as smooth, coordinated skeletal movements, like walking, running, jumping, driving, typing, playing the piano, etc. Patients with cerebellar dysfunction have problems with precise movements, such as walking and balance, and hand and arm movements. The cerebellum looks similar in all animals, from fish to mice to humans. This has been taken as evidence that it performs a common function, such as regulating motor learning and the timing of movements, in all animals. Studies of simple forms of motor learning in the vestibulo–ocular reflex and eye–blink conditioning are demonstrating that timing and amplitude of learned movements are encoded by the cerebellum.

The cerebellum is responsible for coordinating precisely timed activity by integrating motor output with ongoing sensory feedback. It receives extensive projections from sensory–motor areas of the cortex and the periphery and directs it back to premotor and motor cortex [Ghez (1990); Ghez (1991)]. This suggests a role in sensory–motor integration and the timing and execution of human movements. The cerebellum stores patterns of motor control for frequently performed movements, and therefore, its circuits are changed by experience and training. It was termed the adjustable pattern generator in the work of J. Houk and collaborators [Houk et al. (1996)]. Also, it has become the inspiring ‘brain–model’ in the recent robotic research [Schaal and Atkeson (1998); Ivancevic and Ivancevic (2005)].

Comparing the number of its neurons ($10^7$ – $10^8$), to the size of conventional neural networks, suggests that artificial neural nets cannot satisfactorily model the function of this sophisticated ‘super–bio–computer’, as its dimensionality is virtually infinite. Despite a lot of research dedicated to its structure and function (see Houk et al. (1996) and references there cited), the real nature of the cerebellum still remains a ‘mystery’.

The main function of the cerebellum as a motor controller is depicted
Fig. 3.22 Schematic organization of the primary cerebellar circuit. In essence, excitatory inputs, conveyed by collateral axons of Mossy and Climbing fibers activate directly neurons in the Deep cerebellar nuclei. The activity of these latter is also modulated by the inhibitory action of the cerebellar cortex, mediated by the Purkinje cells.

Fig. 3.23 The cerebellum as a motor controller. A coordinated movement is easy to recognize, but we know little about how it is achieved. In search of the neural basis of coordination, a model of spinocerebellar interactions was recently presented in [Apps and Garwicz (2005)], in which the structural and functional organizing principle is a division of the cerebellum into discrete micro–complexes. Each micro–complex is the recipient of a specific motor error signal – that is, a signal that conveys information about an inappropriate movement. These signals are encoded by spinal reflex circuits and conveyed to the cerebellar cortex.
through climbing fibre afferents. This organization reveals salient features of cerebellar information processing, but also highlights the importance of systems level analysis for a fuller understanding of the neural mechanisms that underlie behavior.

3.4.4.3 Hamiltonian Action and Neural Path Integral

Here, we propose a quantum–like adaptive control approach to modelling the ‘cerebellar mystery’. Corresponding to the affine Hamiltonian control function (3.237) we define the affine Hamiltonian control action,

\[ S_{aff}[q,p] = \int_{t_{in}}^{t_{out}} \, dt \, [p_i \dot{q}_i - H_{aff}(q,p)]. \] (3.239)

From the affine Hamiltonian action (3.239) we further derive the associated expression for the neural phase–space path integral (in normal units), representing the cerebellar sensory–motor amplitude \( \langle \text{out} | \text{in} \rangle \),

\[ \langle q^i_{\text{out}}, p^i_{\text{out}} | q^i_{\text{in}}, p^i_{\text{in}} \rangle = \int D[w,q,p] \, e^{i S_{aff}[q,p]} \] (3.240)

\[ = \int D[w,q,p] \exp \left\{ i \int_{t_{in}}^{t_{out}} \, dt \, [p_i \dot{q}_i - H_{aff}(q,p)] \right\}, \]

with

\[ \int D[w,q,p] = \int \prod_{\tau=1}^{n} \frac{w^i(\tau)dp_1(\tau)dq_1(\tau)}{2\pi}, \]

where \( w_i = w_i(t) \) denote the cerebellar synaptic weights positioned along its neural pathways, being continuously updated using the Hebbian–like self–organizing learning rule (3.236). Given the transition amplitude \( \langle \text{out} | \text{in} \rangle \) (3.240), the cerebellar sensory–motor transition probability is defined as its absolute square, \( | \langle \text{out} | \text{in} \rangle |^2 \).

In (3.240), \( q^i_{\text{in}} = q^i_{\text{in}}(t) \), \( q^i_{\text{out}} = q^i_{\text{out}}(t) \); \( p^i_{\text{in}} = p^i_{\text{in}}(t) \), \( p^i_{\text{out}} = p^i_{\text{out}}(t) \); \( t_{in} \leq t \leq t_{out} \), for all discrete time steps, \( t = 1,\ldots,n \to \infty \), and we are allowing for the affine Hamiltonian \( H_{aff}(q,p) \) to depend upon all the \((M \leq N)\) EMA–angles and angular momenta collectively. Here, we actually systematically took a discretized differential time limit of the form \( t_\sigma - t_{\sigma-1} \equiv d\tau \) (both \( \sigma \) and \( \tau \) denote discrete time steps) and wrote \( \frac{q^i_{\sigma} - q^i_{\sigma-1}}{t_\sigma - t_{\sigma-1}} \equiv \dot{q}^i \). For technical details regarding the path integral calculations on Riemannian and symplectic manifolds (including the standard regularization procedures), see Klauder (1997), Klauder (2000).
Now, motor learning occurring in the cerebellum can be observed using functional MR imaging, showing changes in the cerebellar action potential, related to the motor tasks (see, e.g., Mascalchi et al. (2002)). To account for these electro-physiological currents, we need to add the source term $J_i(t)q^i(t)$ to the affine Hamiltonian action (3.239), (the current $J_i = J_i(t)$ acts as a source $J_iA^i$ of the cerebellar electrical potential $A^i = A^i(t)$),

$$S_{aff}[q,p,J] = \int_{t_{in}}^{t_{out}} d\tau \left[p_i \dot{q}^i - H_{aff}(q,p) + J_i q^i\right],$$

which, subsequently gives the cerebellar path integral with the action potential source, coming either from the motor cortex or from other subcortical areas.

Note that the standard Wick rotation: $t \mapsto \frac{i}{t}$ (see Klauder (1997) Klauder (2000)), makes all our path integrals real, i.e.,

$$\int D[w,q,p] e^{i S_{aff}[q,p]} \rightarrow \int D[w,q,p] e^{- S_{aff}[q,p]},$$

while their subsequent discretization gives the standard thermodynamic partition functions,

$$Z = \sum_j e^{-w_j E_j/T},$$

where $E_j$ is the energy eigenvalue corresponding to the affine Hamiltonian $H_{aff}(q,p)$, $T$ is the temperature-like environmental control parameter, and the sum runs over all energy eigenstates (labelled by the index $j$). From (5.16), we can further calculate all statistical and thermodynamic system properties (see Feynman (1972)), as for example, transition entropy $S = k_B \ln Z$, etc.

### 3.4.5 Topological Phase Transitions and Hamiltonian Chaos

#### 3.4.5.1 Phase Transitions in Hamiltonian Systems

Recall that phase transitions (PTs) are phenomena which bring about qualitative physical changes at the macroscopic level in presence of the same microscopic forces acting among the constituents of a system. Their mathematical description requires to translate into quantitative terms the mentioned qualitative changes. The standard way of doing this is to consider...
how the values of thermodynamic observables, get in laboratory experiments, vary with temperature, or volume, or an external field, and then to associate the experimentally observed discontinuities at a PT to the appearance of some kind of singularity entailing a loss of analyticity. Despite the smoothness of the statistical measures, after the Yang–Lee Theorem [Yang and Lee (1952)] we know that in the $N \to \infty$ limit non-analytic behaviors of thermodynamic functions are possible whenever the analyticity radius in the complex fugacity plane shrinks to zero, because this entails the loss of uniform convergence in $N$ (number of degrees of freedom) of any sequence of real-valued thermodynamic functions, and all this depends on the distribution of the zeros of the grand canonical partition function. Also the other developments of the rigorous theory of PTs [Georgii (1988) Ruelle (1978)], identify PTs with the loss of analyticity.

In this subsection we will address a recently proposed geometric approach to thermodynamic phase transitions (see [Caiani et al. (1997); Franzosi et al. (1999) [Franzosi et al. (2000) [Franzosi and Pettini (2004)]]). Given any Hamiltonian system, the configuration space can be equipped with a metric, in order to get a Riemannian geometrization of the dynamics. At the beginning, several numerical and analytical studies of a variety of models showed that the fluctuation of the curvature becomes singular at the transition point. Then the following conjecture was proposed in [Caiani et al. (1997)]: The phase transition is determined by a change in the topology of the configuration space, and the loss of analyticity in the thermodynamic observables is nothing but a consequence of such topological change. The latter conjecture is also known as the topological hypothesis.

The topological hypothesis states that suitable topology changes of equipotential submanifolds of the Hamiltonian system’s configuration manifold can entail thermodynamic phase transitions [Franzosi et al. (2000)]. The authors of the topological hypothesis gave both a theoretical argument and numerical demonstration in case of 2$D$ lattice $\phi^4$ model. They considered classical many-particle (or many-subsystem) systems described by standard mechanical Hamiltonians
\[ H(p, q) = \sum_{i=1}^{N} \frac{p_i^2}{2m} + V(q), \tag{3.242} \]
where the coordinates $q^i = q^i(t)$ and momenta $p_i = p_i(t), \ (i = 1, \ldots, N)$, have continuous values and the system’s potential energy $V(q)$ is bounded below.
Now, assuming a large number of subsystems $N$, the statistical behavior of physical systems described by Hamiltonians of the type (3.242) is usually encompassed, in the system’s canonical ensemble, by the partition function in the system’s phase–space

$$Z_N(\beta) = \int \prod_{i=1}^{N} dp_i dq_i e^{-\beta H(p,q)} = \left(\frac{\pi}{\beta}\right)^{\frac{N}{2}} \int \prod_{i=1}^{N} dq_i e^{-\beta V(q)}$$

$$= \left(\frac{\pi}{\beta}\right)^{\frac{N}{2}} \int_0^\infty dv e^{-\beta v} \int_{M_v} d\sigma / \left\| \nabla V \right\|,$$

(3.243)

where the last term is written using a co–area formula [Federer (1969)], and $v$ labels the equipotential hypersurfaces $M_v$ of the system’s configuration manifold $M$.

Equation (4.118) shows that for Hamiltonians (7.10) the relevant statistical information is contained in the canonical configurational partition function $Z_C^c$.

$$Z_C^c = \int \prod_{i=1}^{N} dq_i \exp[-\beta V(q)].$$

Therefore, partition function $Z_C^c$ is decomposed – in the last term of equation (4.118) – into an infinite summation of geometric integrals, $\int_{M_v} d\sigma / \left\| \nabla V \right\|$, defined on the $\{M_v\}_{v \in \mathbb{R}}$. Once the microscopic interaction potential $V(q)$ is given, the configuration space of the system is automatically foliated into the family $\{M_v\}_{v \in \mathbb{R}}$ of these equipotential hypersurfaces.

Now, from standard statistical mechanical arguments we know that, at any given value of the inverse temperature $\beta$, the larger the number $N$ of particles the closer to $M_v \equiv M_{u_{\beta}}$ are the microstates that significantly contribute to the averages – computed through $Z_N(\beta)$ – of thermodynamic observables. The hypersurface $M_{u_{\beta}}$ is the one associated with the average potential energy computed at a given $\beta$,

$$u_{\beta} = (Z_N^c)^{-1} \int \prod_{i=1}^{N} dq_i V(q) \exp[-\beta V(q)].$$

Thus, at any $\beta$, if $N$ is very large the effective support of the canonical measure shrinks very close to a single $M_v = M_{u_{\beta}}$.

Explicitly, the topological hypothesis reads: the basic origin of a phase transition lies in a suitable topology change of the $\{M_v\}$, occurring at some
This topology change induces the singular behavior of the thermodynamic observables at a phase transition. By change of topology we mean that $\{M_v\}_{v<v_c}$ are not diffeomorphic to the $\{M_v\}_{v>v_c}$. In other words, canonical measure should ‘feel’ a big and sudden change of the topology of the equipotential hypersurfaces of its underlying support, the consequence being the appearance of the typical signals of a phase transition.

This point of view has the interesting consequence that – also at finite $N$ – in principle different mathematical objects, i.e., manifolds of different cohomology type, could be associated to different thermodynamical phases, whereas from the point of view of measure theory [Yang and Lee (1952)] the only mathematical property available to signal the appearance of a phase transition is the loss of analyticity of the grand–canonical and canonical averages, a fact which is compatible with analytic statistical measures only in the mathematical $N \to \infty$ limit.

As it is conjectured that the counterpart of a phase transition is a breaking of diffeomorphicity among the surfaces $M_v$, it is appropriate to choose a diffeomorphism invariant to probe if and how the topology of the $M_v$ changes as a function of $v$. This is a very challenging task because we have to deal with high dimensional manifolds. Fortunately a topological invariant exists whose computation is feasible, yet demands a big effort. Recall (from subsection [Ivancevic and Ivancevic (2007b)]) that this is the Euler characteristic, a diffeomorphism invariant of the system’s configuration manifold, expressing its fundamental topological information.

### 3.4.5.2 Geometry of the Largest Lyapunov Exponent

Now, the topological hypothesis has recently been promoted into a topological Theorem [Franzosi and Pettini (2004)]. The new Theorem says that non–analyticity is the ‘shadow’ of a more fundamental phenomenon occurring in the system’s configuration manifold: a topology change within the family of equipotential hypersurfaces (3.244). This topological approach to PTs stems from the numerical study of the Hamiltonian dynamical counterpart of phase transitions, and precisely from the observation of discontinuous or cuspy patterns, displayed by the largest Lyapunov exponent at the transition energy (or temperature).

Recall that the Lyapunov exponents measure the strength of dynamical chaos and cannot be measured in laboratory experiments, at variance with thermodynamic observables, thus, being genuine dynamical observables they are only measurable in numerical simulations of the microscopic...
dynamics. To get a hold of the reason why the largest Lyapunov exponent \( \lambda_1 \) should probe configuration space topology, let us first remember that for standard Hamiltonian systems, \( \lambda_1 \) is computed by solving the tangent dynamics equation for Hamiltonian systems (see Jacobi equation of geodesic deviation (5.22)),

\[
\ddot{\xi}_i + \left( \frac{\partial^2 V}{\partial q^i \partial q^j} \right)_{q(t)} \xi^j = 0, \tag{3.245}
\]

which, for the nonlinear Hamiltonian system

\[
\begin{align*}
\dot{q}^1 &= p_1, \\
\dot{p}_1 &= -\partial_{q_1} V, \\
&\vdots \\
\dot{q}^N &= p_N, \\
\dot{p}_N &= -\partial_{q_N} V,
\end{align*}
\]

expands into linearized Hamiltonian dynamics

\[
\begin{align*}
\dot{\xi}_1 &= \xi_{N+1}, \\
\dot{\xi}_{N+1} &= -\sum_{j=1}^{N} \left( \frac{\partial^2 V}{\partial q_1 \partial q_j} \right)_{q(t)} \xi_j, \\
&\vdots \\
\dot{\xi}_n &= \xi_{2N}, \\
\dot{\xi}_{2N} &= -\sum_{j=1}^{N} \left( \frac{\partial^2 V}{\partial q_{N} \partial q_j} \right)_{q(t)} \xi_j. \tag{3.246}
\end{align*}
\]

Using (3.245) we can get the analytical expression for the largest Lyapunov exponent

\[
\lambda_1 = \lim_{t \to \infty} \frac{1}{t} \log \left[ \frac{\xi_1^2(t) + \cdots + \xi_N^2(t) + \xi_{1}^2(t) + \cdots + \xi_N^2(t)}{\xi_1^2(0) + \cdots + \xi_N^2(0) + \cdots + \xi_N^2(0)} \right]^{1/2}. \tag{3.247}
\]

If there are critical points of \( V \) in configuration space, that is points \( q_c = [q^1, \ldots, q^N] \) such that \( \nabla V(q)|_{q=q_c} = 0 \), according to the Morse lemma (see e.g., [Hirsch (1976)]), in the neighborhood of any critical point \( q_c \) there always exists a coordinate system \( \tilde{q}(t) = [\tilde{q}^1(t), \ldots, \tilde{q}^N(t)] \) for which

\[
V(\tilde{q}) = V(q_c) - (\tilde{q}^1)^2 - \cdots - (\tilde{q}^k)^2 + (\tilde{q}^{k+1})^2 + \cdots + (\tilde{q}^N)^2, \tag{3.248}
\]

where \( k \) is the index of the critical point, i.e., the number of negative eigenvalues of the Hessian of \( V \). In the neighborhood of a critical point, equation (3.248) yields

\[
\frac{\partial^2 V}{\partial q^i \partial q^j} = \pm \delta_{ij},
\]
which, substituted into equation (3.245), gives \( k \) unstable directions which contribute to the exponential growth of the norm of the tangent vector \( \xi = \xi(t) \). This means that the strength of dynamical chaos, measured by the largest Lyapunov exponent \( \lambda_1 \), is affected by the existence of critical points of \( V \). In particular, let us consider the possibility of a sudden variation, with the potential energy \( v \), of the number of critical points (or of their indexes) in configuration space at some value \( v_c \), it is then reasonable to expect that the pattern of \( \lambda_1(v) \) – as well as that of \( \lambda_1(E) \) since \( v = v(E) \) – will be consequently affected, thus displaying jumps or cusps or other singular patterns at \( v_c \).

On the other hand, recall that Morse theory teaches us that the existence of critical points of \( V \) is associated with topology changes of the hypersurfaces \( \{ M_v \}_{v \in \mathbb{R}} \), provided that \( V \) is a good Morse function (that is: bounded below, with no vanishing eigenvalues of its Hessian matrix). Thus the existence of critical points of the potential \( V \) makes possible a conceptual link between dynamics and configuration space topology, which, on the basis of both direct and indirect evidence for a few particular models, has been formulated as a topological hypothesis about the relevance of topology for PTs phenomena (see [Franzosi et al. (2000); Franzosi and Pettini (2004); Grinza and Mossa (2004)]) in PTs phenomena (see [Franzosi et al. (2000); Franzosi and Pettini (2004); Grinza and Mossa (2004)]).

Here we give two simple examples of standard Hamiltonian systems of the form (3.242), namely Peyrard–Bishop system and mean–field XY model.

**Peyrard–Bishop Hamiltonian System**

The Peyrard–Bishop system [Peyrard and Bishop (1989)] exhibits a second–order phase transition. It is defined by the following potential energy

\[
V(q) = \sum_{i=1}^{N} \left[ \frac{K}{2} (q_{i+1} - q_i)^2 + D(e^{-aq_i} - 1)^2 + Dhaq_i \right],
\]

which represents the energy of a string of \( N \) base pairs of reduced mass \( m \). Each hydrogen bond is characterized by the stretching \( q_i \) and its conjugate momentum \( p_i = m \dot{q}_i \). The elastic transverse force between neighboring pairs is tuned by the constant \( K \), while the energy \( D \) and the inverse

\[24\] The Peyrard–Bishop system has been proposed as a simple model for describing the DNA thermally induced denaturation [Grinza and Mossa (2004)].
length $a$ determine, respectively, the plateau and the narrowness of the on-site potential well that mimics the interaction between bases in each pair. It is understood that $K$, $D$, and $a$ are all positive parameters. The transverse, external stress $h \geq 0$ is a computational tool useful in the evaluation of the susceptibility. Our interest in it lies in the fact that a phase transition can occur only when $h = 0$. We assume periodic boundary conditions.

The transfer operator technique [Dauxois et al. (2002)] maps the problem of computing the classical partition function into the easier task of evaluating the lowest energy eigenvalues of a ‘quantum’ mechanical Morse oscillator (no real quantum mechanics is involved, since the temperature plays the role of $\hbar$). One can then observe that, as the temperature increases, the number of levels belonging to the discrete spectrum decreases, until for some critical temperature $T_c = 2\sqrt{2KD/(ak_B)}$ only the continuous spectrum survives. This passage from a localized ground state to an unnormalizable one corresponds to the second–order phase transition of the statistical model. Various critical exponents can be analytically computed and all applicable scaling laws can be checked. The simplicity of this model permits an analytical computation of the largest Lyapunov exponent by exploiting the geometric method proposed in [Caiani et al. (1997)].

**Mean–Field XY Hamiltonian System**

The mean–field XY model describes a system of $N$ equally coupled planar classical rotators (see [Antoni and Ruffo (1995); Casetti et al. (1999)]). It is defined by a Hamiltonian of the class (3.242) where the potential energy is

$$V(\varphi) = \frac{J}{2N} \sum_{i,j=1}^{N} \left[ 1 - \cos(\varphi_i - \varphi_j) \right] - h \sum_{i=1}^{N} \cos \varphi_i. \quad (3.250)$$

Here $\varphi_i \in [0, 2\pi]$ is the rotation angle of the $i$th rotator and $h$ is an external field. Defining at each site $i$ a classical spin vector $s_i = (\cos \varphi_i, \sin \varphi_i)$ the model describes a planar (XY) Heisenberg system with interactions of equal strength among all the spins. We consider only the ferromagnetic case $J > 0$; for the sake of simplicity, we set $J = 1$. The equilibrium statistical mechanics of this system is exactly described, in the thermodynamic limit, by the mean–field theory [Antoni and Ruffo (1995)]. In the limit $h \to 0$, the system has a continuous phase transition, with classical critical exponents, at $T_c = 1/2$, or $\varepsilon_c = 3/4$, where $\varepsilon = E/N$ is the energy per particle.
The Lyapunov exponent $\lambda_1$ of this system is extremely sensitive to the phase transition. According to reported numerical simulations (see [Casetti et al. (1999)]), $\lambda_1(\varepsilon)$ is positive for $0 < \varepsilon < \varepsilon_c$, shows a sharp maximum immediately below the critical energy, and drops to zero at $\varepsilon_c$ in the thermodynamic limit, where it remains zero in the whole region $\varepsilon > \varepsilon_c$, which corresponds to the thermodynamic disordered phase. In fact in this phase the system is integrable, reducing to an assembly of uncoupled rotators.

3.4.5.3 Euler Characteristics of Hamiltonian Systems

Recall that Euler characteristic $\chi$ is a number that is a characterization of the various classes of geometric figures based only on the topological relationship between the numbers of vertices $V$, edges $E$, and faces $F$, of a geometric Figure. This number, $\chi = F - E + V$, is the same for all figures the boundaries of which are composed of the same number of connected pieces. Therefore, the Euler characteristic is a topological invariant, i.e., any two geometric figures that are homeomorphic to each other have the same Euler characteristic.

More specifically, a standard way to analyze a geometric Figure is to fragment it into other more familiar objects and then to examine how these pieces fit together. Take for example a surface $M$ in the Euclidean 3D space. Slice $M$ into pieces that are curved triangles (this is called a triangulation of the surface). Then count the number $F$ of faces of the triangles, the number $E$ of edges, and the number $V$ of vertices on the tesselated surface. Now, no matter how we triangulate a compact surface $\Sigma$, its Euler characteristic, $\chi(\Sigma) = F - E + V$, will always equal a constant which is characteristic of the surface and which is invariant under diffeomorphisms $\phi : \Sigma \to \Sigma'$.

At higher dimensions this can be again defined by using higher dimensional generalizations of triangles (simplexes) and by defining the Euler characteristic $\chi(M)$ of the nD manifold $M$ to be the alternating sum:

$$
\chi(M) = \sum_{k=0}^{n} (-1)^k \{\text{number of faces of dimension } k\}.
$$

and then define the Euler characteristic of a manifold as the Euler characteristic of any simplicial complex homeomorphic to it. With this definition, circles and squares have Euler characteristic 0 and solid balls have Euler
characteristic 1.

The Euler characteristic $\chi$ of a manifold is closely related to its genus $g$ as $\chi = 2 - 2g$.\footnote{Recall that the genus of a topological space such as a surface is a topologically invariant property defined as the largest number of nonintersecting simple closed curves that can be drawn on the surface without separating it, i.e., an integer representing the maximum number of cuts that can be made through it without rendering it disconnected. This is roughly equivalent to the number of holes in it, or handles on it. For instance: a point, line, and a sphere all have genus 0; a torus has genus 1, as does a coffee cup as a solid object (solid torus), a Möbius strip, and the symbol 0; the symbols 8 and $B$ have genus 2; etc.}

Recall that a more standard topological definition of $\chi(M)$ is

$$\chi(M) = \sum_{k=0}^{n} (-1)^{k} b_k(M), \quad (3.251)$$

where $b_k$ are the $k$th Betti numbers of $M$.

In general, it would be hopeless to try to practically calculate $\chi(M)$ from (3.251) in the case of non-trivial physical models at large dimension. Fortunately, there is a possibility given by the Gauss–Bonnet formula, that relates $\chi(M)$ with the total Gauss–Kronecker curvature of the manifold \cite{Ivancevic and Ivancevic (2007b)},

$$\chi(M) = \gamma \int_M K_G d\sigma, \quad (3.252)$$

which is valid for even dimensional hypersurfaces of Euclidean spaces $\mathbb{R}^N$ [here dim$(M) = n = N - 1$], and where:

$$\gamma = \frac{2}{\text{Vol}(S^n_1)}$$

is twice the inverse of the volume of an $n$–dimensional sphere of unit radius $S^n_1$; $K_G$ is the Gauss–Kronecker curvature of the manifold;

$$d\sigma = \sqrt{\det(g)} \, dx^1 dx^2 \cdots dx^n$$

is the invariant volume measure of $M$ and $g$ is its Riemannian metric (induced from $\mathbb{R}^N$). Let us briefly sketch the meaning and definition of the Gauss–Kronecker curvature. The study of the way in which an $n$–surface $M$ curves around in $\mathbb{R}^N$ is measured by the way the normal direction changes as we move from point to point on the surface. The rate of change of the normal direction $\xi$ at a point $x \in M$ in direction $v$ is described by...
Quantum Leap

the shape operator

\[ L_x(v) = -\mathcal{L}_v \xi = [v, \xi], \]

where \( v \) is a tangent vector at \( x \) and \( \mathcal{L}_v \) is the Lie derivative, hence

\[ L_x(v) = -(\nabla \xi_1 \cdot v, \ldots, \nabla \xi_{n+1} \cdot v); \]

gradients and vectors are represented in \( \mathbb{R}^N \). As \( L_x \) is an operator of the tangent space at \( x \) into itself, there are \( n \) independent eigenvalues \( \kappa_1(x), \ldots, \kappa_n(x) \) which are called the principal curvatures of \( M \) at \( x \) [Thorpe (1979)]. Their product is the Gauss–Kronecker curvature:

\[ K_{G}(x) = \prod_{i=1}^{n} \kappa_i(x) = \det(L_x). \]

Alternatively, recall that according to the Morse theory, it is possible to understand the topology of a given manifold by studying the regular critical points of a smooth Morse function defined on it. In our case, the manifold \( M \) is the configuration space \( \mathbb{R}^N \) and the natural choice for the Morse function is the potential \( V(q) \). Hence, one is lead to define the family \( M_v \) of submanifolds of \( M \).

A full characterization of the topological properties of \( M_v \) generally requires the critical points of \( V(q) \), which means solving the equations

\[ \partial_q V = 0, \quad (i = 1, \ldots, N). \quad (3.253) \]

Moreover, one has to calculate the indexes of all the critical points, that is the number of negative eigenvalues of the Hessian \( \partial^2 V / (\partial q_i \partial q_j) \). Then the Euler characteristic \( \chi(M_v) \) can be computed by means of the formula

\[ \chi(M_v) = \sum_{k=0}^{N} (-1)^k \mu_k(M_v), \quad (3.254) \]

where \( \mu_k(M_v) \) is the total number of critical points of \( V(q) \) on \( M_v \), which have index \( k \), i.e., the so-called Morse numbers of a manifold \( M \), which happen to be upper bounds of the Betti numbers,

\[ b_k(M) \leq \mu_k(M) \quad (k = 0, \ldots, n). \quad (3.255) \]

Among all the Morse functions on a manifold \( M \), there is a special class, called perfect Morse functions, for which the Morse inequalities \( (3.255) \) hold as equalities. Perfect Morse functions characterize completely the topology of a manifold.
Now, we continue with our two examples started before.

**Peyrard–Bishop System.** If applied to any generic model, calculation of (3.254) turns out to be quite formidable, but the exceptional simplicity of the Peyrard–Bishop model (3.249) makes it possible to carry on completely the topological analysis without invoking equation (3.254).

For the potential in exam, equation (3.253) results in the nonlinear system
\[
\frac{a}{R} (q_{i+1} - 2q_i + q_{i-1}) = h - 2(e^{-2aq_i} - e^{-aq_i}),
\]
where \(R = Da^2/K\) is a dimensionless ratio. It is easy to verify that a particular solution is given by
\[
q_i = -\frac{1}{a} \ln \frac{1 + \sqrt{1 + 2h}}{2}, \quad (i = 1, \ldots, N).
\]
The corresponding minimum of potential energy is
\[
V_{\text{min}} = ND \left( \frac{1 + h - \sqrt{1 + 2h}}{2} - h \ln \frac{1 + \sqrt{1 + 2h}}{2} \right).
\]

**Mean–Field XY Model.** In the case of the mean–field XY model (3.250) it is possible to show analytically that a topological change in the configuration space exists and that it can be related to the thermodynamic phase transition. Consider again the family \(M_v\) of submanifolds of the configuration space defined in (3.244); now the potential energy per degree of freedom is that of the mean–field XY model, i.e.,
\[
\mathcal{V}(\varphi) = \frac{V(\varphi)}{N} = \frac{J}{2N^2} \sum_{i,j=1}^{N} \left[ 1 - \cos(\varphi_i - \varphi_j) \right] - h \sum_{i=1}^{N} \cos \varphi_i,
\]
where \(\varphi_i \in [0, 2\pi]\). Such a function can be considered a Morse function on \(M\), so that, according to Morse theory, all these manifolds have the same topology until a critical level \(V^{-1}(v_c)\) is crossed, where the topology of \(M_v\) changes.

A change in the topology of \(M_v\) can only occur when \(v\) passes through a critical value of \(\mathcal{V}\). Thus in order to detect topological changes in \(M_v\) we have to find the critical values of \(\mathcal{V}\), which means solving the equations
\[
\partial_{\varphi_i} \mathcal{V}(\varphi) = 0, \quad (i = 1, \ldots, N).
\]
For a general potential energy function \(\mathcal{V}\), the solution of (3.256) would be a formidable task, but in the case of the mean–field XY model, the mean–
field character of the interaction greatly simplifies the analysis, allowing an
analytical treatment of (3.256); moreover, a projection of the configuration
space onto a 2D plane is possible \cite{Casetti99, Casetti03}. 

_{Quantum Leap_}
Chapter 4

Quantum Universe

In this chapter we apply the methods elaborated so far to present the ‘Holy Grail’ of modern physical and cosmological science, the search for the ‘theory of everything’ and the ‘true’ cosmological dynamics.

4.1 Search for Quantum Gravity

4.1.1 What Is Quantum Gravity?

The landscape of fundamental physics has changed substantially during the last few decades. Not long ago, our understanding of the weak and strong interactions was very confused, while general relativity was almost totally disconnected from the rest of physics and was empirically supported by little more than its three classical tests. Then two things have happened. The \(SU(3) \times SU(2) \times U(1)\) Standard Model has found a dramatic empirical success, showing that quantum field theory (QFT) is capable of describing all accessible fundamental physics, or at least all non-gravitational physics. At the same time, general relativity (GR) has undergone an extraordinary ‘renaissance’, finding widespread application in astrophysics and cosmology, as well as novel vast experimental support – so that today GR is basic physics needed for describing a variety of physical systems we have access to, including advanced technological systems \[\text{Ashby (1997)}\].

These two parallel developments have moved fundamental physics to a position in which it has rarely been in the course of its history: We have today a group of fundamental laws, the Standard Model and GR, which –even if it cannot be regarded as a satisfactory global picture of Nature– is perhaps the best confirmed set of fundamental theories after Newton’s universal gravitation and Maxwell’s electromagnetism. More importantly,
there are not today experimental facts that openly challenge or escape this set of fundamental laws. In this unprecedented state of affairs, a large number of theoretical physicists from different backgrounds have begun to address the piece of the puzzle which is clearly missing: combining the two halves of the picture and understanding the quantum properties of the gravitational field. Equivalently, understanding the quantum properties of space–time. Interest and researches in quantum gravity have thus increased sharply in recent years. And the problem of understanding what is a quantum space–time is today at the core of fundamental physics.

Today we have some well developed and reasonably well defined tentative theories of quantum gravity. String theory and loop quantum gravity are the two major examples. Within these theories definite physical results have been obtained, such as the explicit computation of the ‘quanta of geometry’ and the derivation of the black hole entropy formula. Furthermore, a number of fresh new ideas, like noncommutative geometry, have entered quantum gravity. For an overview of the problem of quantum gravity, see [Isham (1997)].

4.1.2 Main Approaches to Quantum Gravity

4.1.2.1 String Theory

Recall from the previous chapter that string theory is by far the research direction which is presently most investigated. String theory presently exists at two levels. First, there is a well developed set of techniques that define the string perturbation expansion over a given metric background. Second, the understanding of the non–perturbative aspects of the theory has much increased in recent years [Polchinski (1995)] and in the string community there is a widespread faith, supported by numerous indications, in the existence of a yet-to-be-found full non–perturbative theory, capable of generating the perturbation expansion. There are attempts of constructing this non–perturbative theory, generically denoted $M$ theory. The currently popular one is Matrix–theory, of which it is far too early to judge the effectiveness [Matacz (2002); Ishibashi et al. (1997)].

The claim that string theory solves QG is based on two facts. First, the string perturbation expansion includes the graviton. More precisely, one of the string modes is a massless spin two, and helicity $±2$, particle. Such a particle necessarily couples to the energy–momentum tensor of the rest of
the fields [Weinberg (1964); Weinberg (1980)] and gives general relativity to a first approximation. Second, the perturbation expansion is consistent if the background geometry over which the theory is defined satisfies a certain consistency condition; this condition turns out to be a high energy modification of the Einstein’s equations. The hope is that such a consistency condition for the perturbation expansion will emerge as a full–fledged dynamical equation from the yet–to–be–found non–perturbative theory.

From the point of view of the problem of quantum gravity, the relevant physical results from string theory are two [Rovelli (1997)]:

- **Black hole entropy.** The most remarkable physical results for quantum gravity is the derivation of the Bekenstein–Hawking formula for the entropy of a black hole as a function of the horizon area. This beautiful result has been obtained by Strominger and Vafa (1996), and has then been extended in various directions. The result indicates that there is some unexpected internal consistency between string theory and QFT on curved space.

- **Microstructure of space–time.** There are indications that in string theory the space–time continuum is meaningless below the Planck length. An old set of results on very high energy scattering amplitudes indicates that there is no way of probing the space–time geometry at very short distances. What happens is that in order to probe smaller distance one needs higher energy, but at high energy the string ‘opens up from being a particle to being a true string’ which is spread over space–time, and there is no way of focusing a string’s collision within a small space–time region.

More recently, in the non–perturbative formulation of the Matrix–theory [Matacz (2002)], the space–time coordinates of the string $x^i$ are replaced by matrices $(X^i)_{a,b}$. This can perhaps be viewed as a new interpretation of the space–time structure. The continuous space–time manifold emerges only in the long distance region, where these matrices are diagonal and commute; while the space–time appears to have a noncommutative discretized structure in the short distance regime. This features are still poorly understood, but they have intriguing resonances with noncommutative geometry [Connes et al. (1998)] and loop quantum gravity [Rovelli (1998)].

A key difficulty in string theory is the lack of a complete non–perturbative formulation. During the last year, there has been excitement for some tentative non–perturbative formulations [Matacz (2002)]; but it
is far too early to understand if these attempts will be successful. Many
previously highly acclaimed ideas have been rapidly forgotten.

A distinct and even more serious difficulty of string theory is the lack
of a background independent formulation of the theory. In the words of Ed
Witten:

‘Finding the right framework for an intrinsic, background inde-
pendent formulation of string theory is one of the main problems in
string theory, and so far has remained out of reach... This problem
is fundamental because it is here that one really has to address the
question of what kind of geometrical object the string represents.’

Most of string theory is conceived in terms of a theory describing ex-
citations over this or that background, possibly with connections between
different backgrounds. This is also true for (most) non–perturbative for-
mullations such as Matrix theory. For instance, the (bosonic part of the)
Lagrangian of Matrix–theory is

$$L \sim \frac{1}{2} \text{Tr} \left( \dot{X}^2 + \frac{1}{2}[X^i, X^j]^2 \right).$$

The indices that label the matrices $X^i$ are raised and lowered with a
Minkowski metric, and the theory is Lorentz invariant. In other words,
the Lagrangian is really

$$L \sim \frac{1}{2} \text{Tr} \left( g^{00} \dot{X}_i \dot{X}_j + \frac{1}{2} g^{ik} g^{jl} [X_i, X_j][X_k, X_l] \right),$$

where $g$ is the flat metric of the background. This shows that there is a
non–dynamical metric, and an implicit flat background in the action of the
theory.

However, the world is not formed by a fixed background over which
things happen. The background itself is dynamical. In particular, for
instance, the theory should contain quantum states that are quantum su-
perpositions of different backgrounds – and presumably these states play
an essential role in the deep quantum gravitational regime, namely in situ-
ations such as the Big–Bang\footnote{Recall that the Big–Bang is the scientific theory that the universe emerged from a
tremendously dense and hot state about 13.7 billion years ago. The theory is based on
the observations indicating the expansion of space in accord with the Robertson-Walker
model of general relativity, as indicated by the Hubble red–shift of distant galaxies taken
together with the cosmological principle. Extrapolated into the past, these observations
show that the universe has expanded from a state in which all the matter and energy
} or the final phase of black hole evaporation.
The absence of a fixed background in nature (or active diffeomorphism invariance) is the key general lessons we have learned from gravitational theories [Rovelli (1997)].

There has been a burst of recent activity in an outgrowth of string theory denoted string cosmology by [Veneziano (1991)]. The aim of string cosmology is to extract physical consequences from string theory by applying it to the Big–Bang. The idea is to start from a Minkowski flat universe; show that this is unstable and therefore will run away from the flat (false–vacuum) state. The evolution then leads to a cosmological model that starts off in an inflationary phase. This scenario is described using mini–superspace technology, in the context of the low energy theory that emerge as limit of string theory. Thus, first one freezes all the massive modes of the string, then one freezes all massless modes except the zero modes (the spatially constant ones), obtaining a finite dimensional theory, which can be quantized non–perturbatively.

4.1.2.2 Loop Quantum Gravity

The second most popular approach to quantum gravity, and the most popular among relativists, is loop quantum gravity [Rovelli (1998)]. Loop quantum gravity is presently the best developed alternative to string theory. Like strings, it is not far from a complete and consistent theory and it yields a corpus of definite physical predictions, testable in principle, on quantum space–time.

Loop quantum gravity, however, attacks the problem from the opposite direction than string theory. It is a non-perturbative and background independent theory to start with. In other words, it is deeply rooted into the conceptual revolution generated by general relativity. In fact, successes in the universe was at an immense temperature and density. Physicists do not widely agree on what happened before this, although general relativity predicts a gravitational singularity.

The term Big–Bang is used both in a narrow sense to refer to a point in time when the observed expansion of the universe (Hubble’s law) began calculated to be 13.7 billion (1.37 × 1010) years ago (2%) - and in a more general sense to refer to the prevailing cosmological paradigm explaining the origin and expansion of the universe, as well as the composition of primordial matter through nucleosynthesis as predicted by the Alpher–Bethe–Gamow theory [Alpher et al. (1948)]. From this model, George Gamow was able to predict in 1948 the existence of cosmic microwave background radiation (CMB). The CMB was discovered in 1964 and corroborated the Big Bang theory, giving it more credence over its chief rival, the steady state theory. For details, see section on Cosmological Dynamics below.
Quantum Leap

and problems of loop quantum gravity are complementary to successes and problems of strings. Loop quantum gravity is successful in providing a consistent mathematical and physical picture of non-perturbative quantum space-time; but the connection to the low energy dynamics is not yet completely clear.

The general idea on which loop quantum gravity is based is the following. The core of quantum mechanics is not identified with the structure of (conventional) QFT, because conventional QFT presupposes a background metric space-time, and is therefore immediately in conflict with GR. Rather, it is identified with the general structure common to all quantum systems. The core of GR is identified with the absence of a fixed observable background space-time structure, namely with active diffeomorphism invariance. Loop quantum gravity is thus a quantum theory in the conventional sense: a Hilbert space and a set of quantum (field) operators, with the requirement that its classical limit is GR with its conventional matter couplings. But it is not a QFT over a metric manifold. Rather, it is a ‘quantum field theory on a differentiable manifold’, respecting the manifold’s invariances and where only coordinate independent quantities are physical.

Technically, loop quantum gravity is based on two inputs [Rovelli (1998); Rovelli (1997)]:

- The formulation of classical GR based on the Ashtekar connection [Ashtekar (1986); Ashtekar (1987); Ashtekar (1991)]. The version of the connection now most popular is not the original complex one, but an evolution of the same, in which the connection is real.
- The choice of the holonomies of this connection, denoted loop variables, as basic variables for the quantum gravitational field [Rovelli and Smolin (1988)].

This second choice determines the peculiar kind of quantum theory being built. Physically, it corresponds to the assumption that excitations with support on a loop are normalizable states. This is the key technical assumption on which everything relies.

It is important to notice that this assumption fails in conventional 4D Yang–Mills theory, because loop-like excitations on a metric manifold are too singular: the field needs to be smeared in more dimensions [Rovelli (1997)]. Equivalently, the linear closure of the loop states is a ‘far too big’ non-separable state space. This fact is the major source of some particle physicists’s suspicion at loop quantum gravity. What makes GR different
Quantum Universe

from 4D Yang–Mills theory, however, is non–perturbative diffeomorphism invariance. The gauge invariant states, in fact, are not localized at all – they are, pictorially speaking, smeared by the (gauge) diffeomorphism group all over the coordinates manifold. More precisely, factoring away the diffeomorphism group takes us down from the state space of the loop excitations, which is ‘too big’, to a separable physical state space of the right size. Thus, the consistency of the loop construction relies heavily on diffeomorphism invariance. In other words, the diff–invariant invariant loop states (more precisely, the diff–invariant spin network states) are not physical excitations of a field on space–time. They are excitations of space–time itself.

Loop quantum gravity was briefly described by [Rovelli (1997)] as follows:

Definition of theory. The mathematical structure of the theory has been put on a very solid basis. Early difficulties have been overcome. In particular, there were three major problems in the theory: the lack of a well defined scalar product, the overcompleteness of the loop basis, and the difficulty of treating the reality conditions.

• The problem of the lack of a scalar product on the Hilbert space has been solved with the definition of a diffeomorphism invariant measure on a space of connections [Ashtekar and Lewandowski (1995)]. Later, it has also became clear that the same scalar product can be defined in a purely algebraic manner [DePietri and Rovelli (1996)]. The state space of the theory is therefore a genuine Hilbert space $\mathcal{H}$.

• The overcompleteness of the loop basis has been solved by the introduction of the spin network states [Rovelli and Smolin (1995)]. A spin network is a graph carrying labels (related to $SU(2)$ representations and called ‘colors’) on its links and its nodes. Each spin network defines a spin network state, and the spin network states form a (genuine, non-overcomplete) orthonormal basis in $\mathcal{H}$.

• The difficulties with the reality conditions have been circumvented by the use of the real formulation [Barbero (1994); Barbero (1995a); Barbero (1995b); Thiemann (1996)]. The kinematics of loop quantum gravity is now defined with a level of rigor characteristic of mathematical physics [Ashtekar and Isham (1992); Ashtekar et al. (1995)] and the theory can be de-
Hamiltonian constraint. A rigorous definition version of the Hamiltonian constraint equation has been constructed. This is anomaly free, in the sense that the constraints algebra closes (but see later on). The Hamiltonian has the crucial properties of acting on nodes only, which implies that its action is naturally discrete and combinatorial [Rovelli and Smolin (1988); Rovelli and Smolin (1994)]. This fact is at the roots of the existence of exact solutions [Rovelli and Smolin (1988)], and of the possible finiteness of the theory.

Matter. The old hope that QFT divergences could be cured by QG has recently received an interesting corroboration. The matter part of the Hamiltonian constraint is well–defined without need of renormalization. Thus, a main possible stumbling block is over: infinities did not appear in a place where they could very well have appeared [Rovelli (1997)].

Black hole entropy. The first important physical result in loop quantum gravity is a computation of black hole entropy [Krasnov (1997); Rovelli (1996a); Rovelli (1996b)].

Quanta of geometry. A very exciting development in quantum gravity in the last years has been by the computations of the quanta of geometry. That is, the computation of the discrete eigenvalues of area and volume.

In quantum gravity, any quantity that depends on the metric becomes an operator. In particular, so do the area $A$ of a given (physically defined) surface, or the volume $V$ of a given (physically defined) spatial region. In loop quantum gravity, these operators can be written explicitly. They are mathematically well defined self–adjoint operators in the Hilbert space $H$. We know from quantum mechanics that certain physical quantities are quantized, and that we can compute their discrete values by computing the eigenvalues of the corresponding operator. Therefore, if we can compute the eigenvalues of the area and volume operators, we have a physical prediction on the possible quantized values that these quantities can take, at the Planck scale. These eigenvalues have been computed in loop quantum gravity. Here is for instance the main sequence of the spectrum of the area

$$A_{\vec{j}} = 8\pi\gamma\hbar G \sum_{i} \sqrt{j_i(j_i + 1)}.$$  \hspace{1cm} (4.3)

$\vec{j} = (j_1, \ldots, j_n)$ is an $n$–tuplet of half–integers, labeling the eigenvalues, $G$ and $\hbar$ are the Newton and Planck constants, and $\gamma$ is a dimensionless
free parameter, denoted the so-called Immirzi parameter [Immirzi (1997)], not determined by the theory. A similar result holds for the volume. The spectrum (4.3) has been rederived and completed using various different techniques [DePietri and Rovelli (1996)]. These spectra represent solid results of loop quantum gravity. Under certain additional assumptions on the behavior of area and volume operators in the presence of matter, these results can be interpreted as a corpus of detailed quantitative predictions on hypothetical Planck scale observations.

Besides its direct relevance, the quantization of the area and the volume is of interest because it provides a physical picture of quantum space–time. The states of the spin network basis are eigenstates of some area and volume operators. We can say that a spin network carries quanta of area along its links, and quanta of volume at its nodes. The magnitude of these quanta is determined by the coloring. For instance, the half–integers \( j_1 \ldots j_n \) in (4.3) are the coloring of the spin network’s links that cross the given surface. Thus, a quantum space–time can be decomposed in a basis of states that can be visualized as made by quanta of volume (the intersections) separated by quanta of area (the links). More precisely, we can view a spin network as sitting on the dual of a cellular decomposition of physical space. The nodes of the spin network sit in the center of the 3–cells, and their coloring determines the (quantized) 3–cell’s volume. The links of the spin network cut the faces of the cellular decomposition, and their color \( \vec{j} \) determine the (quantized) areas of these faces via equation (4.3).

4.1.3 Traditional Approaches to Quantum Gravity

4.1.3.1 Discrete Approaches

Discrete quantum gravity is the program of regularizing classical GR in terms of some lattice theory, quantize this lattice theory, and then study an appropriate continuum limit, as one may do in QCD. There are three main ways of discretizing GR.

Regge Calculus

Regge introduced the idea of triangulating space–time by means of a simplicial complex and using the lengths \( l_i \) of the links of the complex as gravitational variables [Regge (1961)]. The theory can then be quantized by integrating over the lengths \( l_i \) of the links. For a recent review and extensive references see [Williams and Tuckey (1992)]. More recent work
has focused in problems such as the geometry of Regge superspace [Hartle et al. (1997)] and choice of the integration measure.

**Dynamical Triangulations**

Alternatively, one can keep the length of the links fixed, and capture the geometry by means of the way in which the simplices are glued together, namely by the triangulation. The *Einstein–Hilbert action* of Euclidean gravity is approximated by a simple function of the total number of simplices and links, and the theory can be quantized summing over distinct triangulations (for a detailed introduction, see [Ambjørn et al. (1998)]). There are two coupling constants in the theory, roughly corresponding to the Newton and cosmological constants. These define a two dimensional space of theories. The theory has a nontrivial continuum limit if in this parameter space there is a critical point corresponding to a second order phase transition. The theory has phase transition and a critical point. The transition separates a phase with crumpled space–times from a phase with ‘elongated’ spaces which are effectively 2D, with characteristic of a branched polymer [Bakker and Smit (1995); Ambjørn et al. (2001c)]. This polymer structure is surprisingly the same as the one that emerges from loop quantum gravity at short scale. Near the transition, the model appears to produce ‘classical’ $S^4$ space–times, and there is evidence for scaling, suggesting a continuum behavior.

**State Sum Models**

A third road for discretizing GR was opened by a celebrated paper by [Ponzano and Regge (1968)]. They started from a Regge discretization of 3D GR and introduced a second discretization, by posing the so–called *Ponzano–Regge ansatz* that the lengths $l$ assigned to the links are discretized as well, in half–integers in Planck units

$$l = hG \frac{j}{2}, \quad j = 0, 1, \ldots$$  \hspace{1cm} (4.4)\

(Planck length is $hG$ in 3D.) The half integers $j$ associated to the links are denoted ‘coloring’ of the triangulation. Coloring can be viewed as the assignment of a $SU(2)$ irreducible representation to each link of the Regge triangulation. The elementary cells of the triangulation are tetrahedra, which have six links, colored with six $SU(2)$ representations. $SU(2)$ representation theory naturally assigns a number to a sextuplet of representations:
the Wigner $6 - j$ symbol. Rather magically, the product over all tetrahedra of these $6 - j$ symbols converges to (the real part of the exponent of) the Einstein–Hilbert action. Thus, Ponzano and Regge were led to propose a quantization of 3D GR based on the partition function

$$Z \sim \sum_{\text{coloring tetrahedra}} \prod_{\text{tetrahedra}} 6 - j(\text{color of the tetrahedron}), \quad (4.5)$$

where we have neglected some coefficients for simplicity. They also provided arguments indicating that this sum is independent from the triangulation of the manifold.

The formula (4.5) is simple and elegant, and the idea has recently had many surprising and interesting developments. 3D GR was quantized as a topological field theory by Ed Witten in [Witten (1988a)] and using loop quantum gravity in [Ashtekar et al. (1989)]. The Ponzano–Regge quantization based on equation (4.5) was shown to be essentially equivalent to the TQFT quantization in [Ooguri (1992)], and to the loop quantum gravity in [Rovelli (1993)] (for an extensive discussion of 3D quantum gravity, see [Carlip et al. (1995)]).

It turns out that the Ponzano–Regge ansatz (4.4) can be derived from loop quantum gravity [Rovelli (1993)]. Indeed, (4.4) is the 2D version of the 3D formula (4.3), which gives the quantization of the area. Therefore, a key result of quantum gravity of the last years, namely the quantization of the geometry, derived in the loop formalism from a full fledged non–perturbative quantization of GR, was anticipated as an ansatz by the intuition of Ponzano and Regge.

4.1.3.2 Hawking’s Euclidean Quantum Gravity

Hawking’s Euclidean quantum gravity is the approach based on his formal sum over Euclidean geometries (i.e., an Euclidean path integral)

$$Z \sim N \int \mathcal{D}[g] e^{-\int d^4x \sqrt{\mathcal{R}[g]}}. \quad (4.6)$$

Firstly, Hawking’s picture of quantum gravity as a sum–over–space–times, continues to provide a powerful intuitive reference point for most of the research related to quantum gravity. Indeed, many approaches can be seen as attempts to replace the ill–defined and non–renormalizable formal integral (4.6) with a well defined expression. The dynamical triangulation approach (see above) and the spin foam approach (see below) are examples
of attempts to realize Hawking’s intuition. Influence of Euclidean quantum gravity can also be found in the Atiyah axioms for TQFT. Secondly, this approach can be used as an approximate method for describing certain regimes of non-perturbative quantum space–time physics, even if the fundamental dynamics is given by a more complete theory. In this spirit, Hawking and collaborators have continued the investigation of phenomena such as, for instance, pair creation of black holes in a background de Sitter space–time.

4.1.3.3 Effective Perturbative Quantum Gravity

If we expand classical GR around, say, the Minkowski metric,
\[ g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu}(x), \]
and construct a conventional QFT for the field \( h_{\mu\nu}(x) \), we get, as it is well known, a non renormalizable theory. A small but intriguing group of papers has recently appeared, based on the proposal of treating this perturbative theory seriously, as a respectable low energy effective theory by its own. This cannot solve the deep problem of understanding the world in general relativistic quantum terms. But it can still be used for studying quantum properties of space–time in some regimes. This view has been advocated in a convincing way by John Donoghue, who has developed effective field theory methods for extracting physics from non renormalizable quantum GR [Donoghue (1996)].

4.1.3.4 QFT in Curved Space–Time

Quantum field theory in curved space–time is by now a reasonably established theory (see, e.g., Wald (1994) Birrel et al. (1982) Fulling (1989)), predicting physical phenomena of remarkable interest such as particle creation, vacuum polarization effects and Hawking’s black-hole radiance (Hawking (1975)). To be sure, there is no direct nor indirect experimental observation of any of these phenomena, but the theory is quite credible as an approximate theory, and many theorists in different fields would probably agree that these predicted phenomena are likely to be real.

The most natural and general formulation of the theory is within the algebraic approach [Haag (1992)], in which the primary objects are the local observables and the states of interest may all be treated on equal footing (as positive linear functionals on the algebra of local observables), even if they do not belong to the same Hilbert space.
The great merit of QFT on curved space–time is that it has provided us with some very important lessons. The key lesson is that in general one loses the notion of a single preferred quantum state that could be regarded as the ‘vacuum’; and that the concept of ‘particle’ becomes vague and/or observer-dependent in a gravitational context. In a gravitational context, vacuum and particle are necessarily ill defined or approximate concepts. It is perhaps regrettable that this important lesson has not been yet absorbed by many scientists working in fundamental theoretical physics [Rovelli (1997)].

4.1.3.5 *New Approaches to Quantum Gravity*

**Noncommutative Geometry**

Noncommutative geometry is a research program in mathematics and physics which has recently received wide attention and raised much excitement. The program is based on the idea that space–time may have a noncommutative structure at the Planck scale. A main driving force of this program is the radical, volcanic and extraordinary sequence of ideas of A. Connes [Connes (1994)]. Connes observes that what we know about the structure of space–time derives from our knowledge of the fundamental interactions: special relativity derives from a careful analysis of Maxwell theory; Newtonian space–time and general relativity, derived both from a careful analysis of the gravitational interaction. Recently, we have learned to describe weak and strong interactions in terms of the $SU(3) \times SU(2) \times U(1)$ Standard Model. Connes suggests that the Standard Model might hide information on the minute structure of space–time as well. By making the hypothesis that the Standard Model symmetries reflect the symmetry of a noncommutative microstructure of space–time, Connes and Lott are able to construct an exceptionally simple and beautiful version of the Standard Model itself, with the impressive result that the Higgs field appears automatically, as the components of the Yang–Mills connection in the internal ‘noncommutative’ direction [Connes and Lott (1990)]. The theory admits a natural extension in which the space–time metric, or the gravitational field, is dynamical, leading to GR [Chamseddine and Connes (1996)].

The key idea behind a non-commutative space–time is to use algebra instead of geometry in order to describe spaces. Consider a topological (Hausdorff) space $M$. Consider all continuous functions $f$ on $M$. These form an algebra $A$, because they can be multiplied and summed, and the
algebra is commutative. According to a celebrated result, due to Gel’fand, knowledge of the algebra $A$ is equivalent to knowledge of the space $M$, i.e., $M$ can be reconstructed from $A$. In particular, the points $x$ of the manifold can be obtained as the 1D irreducible representations $x$ of $A$, which are all of the form $x(f) = f(x)$. Thus, we can use the algebra of the functions, instead of using the space. In a sense, notices Connes, the algebra is more physical, because we never deal with space–time: we deal with fields, or coordinates, over space–time. But one can capture Riemannian geometry as well, algebraically. Consider the Hilbert space $H$ formed by all the spinor fields on a given Riemannian (spin) manifold. Let $D$ be the (curved) Dirac operator, acting on $H$. We can view $A$ as an algebra of (multiplicative) operators on $H$. Now, from the triple $(H, A, D)$, which Connes calls ‘spectral triple’, one can reconstruct the Riemannian manifold. In particular, it is not difficult to see that the distance between two points $x$ and $y$ can be obtained from these data by
\[
d(x, y) = \sup\{f \in A, \|D.f\| < 1\} |x(f) - y(f)|,
\]
a beautiful surprising algebraic definition of distance. A non-commutative space–time is the idea of describing space–time by a spectral triple in which the algebra $A$ is a non-commutative algebra.

Remarkably, the gravitational field is captured, together with the Yang–Mills field, and the Higgs fields, by a suitable Dirac operator $D$ [Chamseddine and Connes (1996)], and the full action is given simply by the trace of a very simple function of the Dirac operator.

Even if we disregard noncommutativity and the Standard Model, the above construction represents an intriguing re–formulation of conventional GR, in which the geometry is described by the Dirac operator instead than the metric tensor. This formulation has been explored in [Landi (1998)], where it is noticed that the eigenvalues of the Dirac operator are diffeomorphism invariant functions of the geometry, and therefore represent true observables in Euclidean GR. Their Poisson bracket algebra can be explicitly computed in terms of the energy–momentum eigenspinors. Surprisingly, the Einstein equations turn out to be captured by the requirement that the energy momentum of the eigen–spinors scale linearly with the eigenvalues.

Variants of Connes’s version of the idea of non commutative geometry and noncommutative coordinates have been explored by many authors (see, e.g., [Doplicher et al. (1994)]) and intriguing connections with string theory have been suggested [Connes et al. (1998)] [Fröhlich and Gawedzki (1994)].
Null–Surface Formulation

A second new set of ideas comes from [Frittelli et al. (1995)]. These authors have discovered that the (conformal) information about the geometry is captured by suitable families of null hypersurfaces in space–time, and have been able to reformulate GR as a theory of self–interacting families of surfaces. A remarkable aspect of the theory is that physical information about the space–time interior is transferred to null infinity, along null geodesics. Thus, the space–time interior is described in terms of how we would (literally) ‘see it’ from outside. This description is diffeomorphism invariant, and addresses directly the relational localization characteristic of GR: the space–time location of a region is determined dynamically by the gravitational field and is captured by when and where we see the space–time region from infinity. This idea may lead to interesting and physically relevant diffeomorphism invariant observables in quantum gravity. A discussion of the quantum gravitational fuzziness of the space–time points determined by this perspective can be found in [Frittelli et al. (1997)].

Spin Foam Models

From the mathematical point of view, the problem of quantum gravity is to understand what is QFT on a differentiable manifold without metric. A class of well understood QFT’s on manifolds exists. These are the topological quantum field theories (TQFT). Topological field theories are particularly simple field theories. They have as many fields as gauges and therefore no local degree of freedom, but only a finite number of global degrees of freedom. An example is GR in 3D, say on a torus (the theory is equivalent to a Chern–Simons theory). In 3D, the Einstein equations require that the geometry is flat, so there are no gravitational waves. Nevertheless, a careful analysis reveals that the radii of the torus are dynamical variables, governed by the theory. Witten has noticed that theories of this kind give rise to interesting quantum models [Witten (1988b)], and [Atiyah (1989)] has provided a beautiful axiomatic definition of a TQFT. Concrete examples of TQFT have been constructed using Hamiltonian, combinatorial and path integral methods. The relevance of TQFT for quantum gravity has been suggested by many and the recent developments have confirmed these suggestions.

Recall that TQFT is a diffeomorphism invariant QFT. Sometimes, the expression TQFT is used to indicate all diffeomorphism invariant QFT’s.
This has lead to a widespread, but incorrect belief that any diffeomorphism invariant QFT has a finite number of degrees of freedom, unless the invariance is somehow broken, for instance dynamically. This belief is wrong. The problem of quantum gravity is precisely to define a diffeomorphism invariant QFT having an infinite number of degrees of freedom and 'local' excitations. Locality in a gravity theory, however, is different from locality in conventional field theory. This point is often source of confusion. Here is Rovelli’s clarification [Rovelli (1997)]:

- In a conventional field theory on a metric space, the degrees of freedom are local in the sense that they can be localized on the metric manifold (an electromagnetic wave is here or there in Minkowski space).
- In a diffeomorphism invariant field theory such as general relativity, the degrees of freedom are still local (gravitational waves exist), but they are not localized with respect to the manifold. They are nevertheless localized with respect to each other (a gravity wave is three meters apart from another gravity wave, or from a black hole).
- In a topological field theory, the degrees of freedom are not localized at all: they are global, and in finite number (the radius of a torus is not in a particular position on the torus).

The first TQFT directly related to quantum gravity was defined by [Turaev and Viro (1992)]. The Turaev–Viro model is a mathematically rigorous version of the 3D Ponzano-Regge quantum gravity model described above. In the Turaev–Viro theory, the sum (4.5) is made finite by replacing $SU(2)$ with quantum $SU(2)_q$ (with a suitable $q$). Since $SU(2)_q$ has a finite number if irreducible representations, this trick, suggested by [Ooguri (1992); Ooguri (1992)], makes the sum finite. The extension of this model to four dimensions has been actively searched for a while and has finally been constructed by [Crane and Yetter (1993)], again following Ooguri’s ideas. The Crane–Yetter (CY) model is the first example of 4D TQFT. It is defined on a simplicial decomposition of the manifold. The variables are spins ('colors') attached to faces and tetrahedra of the simplicial complex. Each 4–simplex contains 10 faces and 5 tetrahedra, and therefore there are 15 spins associated to it. The action is defined in terms of the quantum Wigner $15 – j$ symbols, in the same manner in which the Ponzano–Regge action is constructed in terms of products of $6 – j$ symbols.

$$Z \sim \sum_{\text{coloring}} \prod_{4\text{-simplices}} 15 – j(\text{color of the 4} \text{- simplex}), \quad (4.8)$$
Quantum Universe (where we have disregarded various factors for simplicity). Crane and Yetter introduced their model independently from loop quantum gravity. However, recall that loop quantum gravity suggests that in 4 dimensions the naturally discrete geometrical quantities are area and volume, and that it is natural to extend the Ponzano–Regge model to 4D by assigning colors to faces and tetrahedra.

The CY model is not a quantization of 4D GR, nor could it be, being a TQFT in strict sense. Rather, it can be formally derived as a quantization of SU(2) BF theory. BF theory is a topological field theory with two fields, a connection $A$, with curvature $F$, and a 2–form $B$ [Horowitz (1989)], with action

$$ S[A, B] = \int B \wedge F. \quad (4.9) $$

However, there is a strict relation between GR and BF. If we add to $SO(3,1)$ BF theory the constraint that the 2–form $B$ is the product of two tetrad 1–forms

$$ B = E \wedge E, \quad (4.10) $$

we get precisely GR. This observation has lead many to suggest that a quantum theory of gravity could be constructed by a suitable modification of quantum BF theory [Baez (1996c)]. This suggestion has become very plausible, with the following construction of the spin foam models.

The key step in development of the spin foam models was taken by [Barbieri (1997)], studying the ‘quantum geometry’ of the simplices that play a role in loop quantum gravity. Barbieri discovered a simple relation between the quantum operators representing the areas of the faces of the tetrahedra. This relation turns out to be the quantum version of the constraint (4.10), which turns BF theory into GR. Barret and Crane [Barret and Crane (1997)] added the Barbieri relation to (the $SO(3,1)$ version of) the CY model. This is equivalent to replacing the 15–j Wigner symbol, with a different function $A_{BC}$ of the colors of the 4–simplex. This replacement defines a ‘modified TQFT’, which has a chance of having general relativity as its classical limit.

The Barret–Crane model is not a TQFT in strict sense. In particular, it is not independent from the triangulation. Thus, a continuum theory has to be formally defined by some suitable sum over triangulations

$$ Z \sim \sum_{\text{triang}} \sum_{\text{coloring}} \prod_{4\text{--simplices}} A_{BC}(\text{color of the } 4\text{--simplex}). \quad (4.11) $$
This essential aspect of the construction, however, is not yet understood.

The Barret Crane model can virtually be obtained also from loop quantum gravity. This is an unexpected convergence of two very different lines of research. Loop quantum gravity is formulated canonically in the frozen time formalism. While the frozen time formalism is in principle complete, in practice it is cumbersome, and anti-intuitive. Our intuition is four dimensional, not three dimensional. An old problem in loop quantum gravity has been to derive a space–time version of the theory. A space–time formulation of quantum mechanics is provided by the sum over histories. A sum over histories can be derived from the Hamiltonian formalism, as Feynman did originally. Loop quantum gravity provides a mathematically well defined Hamiltonian formalism, and one can therefore follow Feynman steps and construct a sum over histories quantum gravity starting from the loop formalism. This has been done in [Reisenberger and Rovelli (1997)]. The sum over histories turns out to have the form of a sum over surfaces.

More precisely, the transition amplitude between two spin network states turns out to be given by a sum of terms, where each term can be represented by a (2D) branched ‘colored’ surface in space–time. A branched colored surface is formed by elementary surface elements carrying a label, that meet on edges, also carrying a labelled; edges, in turn meet in vertices (or branching points, see Figure 4.1). The contribution of one such surfaces to the sum over histories is the product of one term per each branching point of the surface. The branching points represent the ‘vertices’ of this theory, in the sense of Feynman. The contribution of each vertex can be computed algebraically from the ‘colors’ (half integers) of the adjacent surface elements and edges. Thus, space–time loop quantum gravity is defined by the partition function

$$Z \sim \sum_{\text{surfaces}} \sum_{\text{colorings}} \prod_{\text{vertices}} A_{\text{loop}}(\text{color of the vertex}).$$  \hspace{1cm} (4.12)

The vertex $A_{\text{loop}}$ is determined by a matrix elements of the Hamiltonian constraint. The fact that one gets a sum over surfaces is not too surprising, since the time evolution of a loop is a surface. Indeed, the time evolution of a spin network (with colors on links and nodes) is a surface (with colors on surface elements and edges) and the Hamiltonian constraint generates branching points in the same manner in which conventional Hamiltonians generate the vertices of the Feynman diagrams.

Now, (4.12) has the same structure of the Barret–Crane model (4.8). To see this, simply notice that we can view each branched colored surface as
Fig. 4.1 A branched surface with two vertices.

located on the lattice dual to a triangulation. Then each vertex correspond to a 4-simplex; the coloring of the two models matches exactly (elementary surfaces → faces, edges → tetrahedra); and summing over surfaces corresponds to summing over triangulations. The main difference is the different weight at the vertices. The Barret–Crane vertex $A_{BC}$ can be read as a covariant definition a Hamiltonian constraint in loop quantum gravity.

Thus, the space–time formulation of loop quantum GR is a simple modification of a TQFT. This approach provides a 4D pictorial intuition of quantum space–time, analogous to the Feynman graphs description of quantum field dynamics. John Baez has introduced the term ‘spin foam’ for the branched colored surfaces of the model, in honor of John Wheeler’s intuitions on the quantum microstructure of space–time. Spin foams are a precise mathematical implementation of Wheeler’s ‘space–time foam’ suggestions.

4.1.3.6 Hawking’s Black Hole Entropy

A focal point of the research in quantum gravity in the last years has been the discussion of black hole (BH) entropy. This problem has been discussed from a large variety of perspectives and within many different
research programs.

Let us very briefly recall the origin of the problem. In classical GR, future event horizons behave in a manner that has a peculiar thermodynamical flavor. This remark, together with a detailed physical analysis of the behavior of hot matter in the vicinity of horizons, prompted Bekenstein to suggest that there is entropy associated to every horizon. The suggestion was first considered ridicule, because it implies that a black hole is hot and radiates. But then S. Hawking, in a celebrated work [Hawking (1975)], showed that QFT in curved space–time predicts that a black hole emits thermal radiation, precisely at the temperature predicted by Bekenstein, and Bekenstein’s courageous suggestion was fully vindicated. Since then, the entropy of a BH has been indirectly computed in a surprising variety of manners, to the point that BH entropy and BH radiance are now considered almost an established fact by the community, although, of course, they were never observed nor, presumably, they are going to be observed soon. This confidence, perhaps a bit surprising to outsiders, is related to the fact thermodynamics is powerful in indicating general properties of systems, even if we do not control its microphysics. Many hope that the Bekenstein–Hawking radiation could play for quantum gravity a role analogous to the role played by the black body radiation for quantum mechanics. Thus, indirect arguments indicate that a Schwarzschild BH has an entropy

$$S = \frac{1}{4} \frac{A}{\hbar G}$$  \hspace{1cm} (4.13)

The remaining challenge is to derive this formula from first principles [Rovelli (1997)].

Later in the book we will continue our exposition of various approaches to quantum gravity.

4.2 Loop Quantum Gravity

4.2.1 Introduction to Loop Quantum Gravity

Recall (from subsection 4.1 above) that C. Rovelli developed (in the last decade of the 20th Century) the so-called loop approach to quantum gravity (see [Rovelli (1998)] and references therein). The first announcement of this approach was given in [Rovelli and Smolin (1987)]. Together with string theory, this approach provides another serious candidate theory of
quantum gravity. It provides a physical picture of Planck scale quantum geometry, calculation techniques, definite quantitative predictions, and a tool for discussing classical problems such as black hole thermodynamics.

String theory and loop quantum gravity differ not only because they explore distinct physical hypotheses, but also because they are expressions of two separate communities of scientists, which have sharply distinct prejudices, and view the problem of quantum gravity in surprisingly different manners. As Rovelli says: “I heard the following criticism to loop quantum gravity: ‘Loop quantum gravity is certainly physically wrong, because:

(1) it is not supersymmetric, and
(2) is formulated in four dimensions’. But experimentally, the world still insists on looking four–dimensional and not supersymmetric. In my opinion, people should be careful of not being blinded by their own speculation, and mistaken interesting hypotheses (such as supersymmetry and high–dimensions) for established truth. But string theory may claim extremely remarkable theoretical successes and is today the leading and most widely investigated candidate theory of quantum gravity” [Rovelli (1998)].

High energy physics has obtained spectacular successes during this Century, culminated with the (far from linear) establishment of quantum field theory as the general form of dynamics and with the comprehensive success of the $SU(3) \times SU(2) \times U(1)$ Standard Model. Thanks to this success, now a few decades old, physics is in a condition in which it has been very rarely: there are no experimental results that clearly challenge, or clearly escape, the present fundamental theory of the world. The theory we have encompasses virtually everything – except gravitational phenomena. From the point of view of a particle physicist, gravity is then simply the last and weakest of the interactions. It is natural to try to understand its quantum properties using the strategy that has been so successful for the rest of microphysics, or variants of this strategy. The search for a conventional quantum field theory capable of embracing gravity has spanned several decades and, through an adventurous sequence of twists, moments of excitement and disappointments, has lead to string theory. The foundations of string theory are not yet well understood; and it is not yet entirely clear how a supersymmetric theory in 10 or 11 dimensions can be concretely used for deriving comprehensive univocal predictions about our world.

In string theory, gravity is just one of the excitations of a string (or other extended object) living over some background metric space. The existence of such background metric space, over which the theory is defined,
Quantum Leap

is needed for the formulation and for the interpretation of the theory, not only in perturbative string theory, but in the recent attempts of a non-perturbative definition of the theory, such as $M$ theory, as well, in my understanding. Thus, for a physicist with a high energy background, the problem of quantum gravity is now reduced to an aspect of the problem of understanding what is the mysterious non-perturbative theory that has perturbative string theory as its perturbation expansion, and how to extract information on Planck scale physics from it.

For a relativist, on the other hand, the idea of a fundamental description of gravity in terms of physical excitations over a background metric space sounds physically very wrong. The key lesson learned from general relativity is that there is no background metric over which physics happens. The world is more complicated than that. Indeed, for a relativist, general relativity is much more than the field theory of a particular force. Rather, it is the discovery that certain classical notions about space and time are inadequate at the fundamental level; they require modifications which are possibly as basics as the ones that quantum mechanics introduced. One of such inadequate notions is precisely the notion of a background metric space (flat or curved), over which physics happens. This profound conceptual shift has led to the understanding of relativistic gravity, to the discovery of black holes, to relativistic astrophysics and to modern cosmology.

From Newton to the beginning of this Century, physics has had a solid foundation in a small number of key notions such as space, time, causality and matter. In spite of substantial evolution, these notions remained rather stable and self-consistent. In the first quarter of this Century, quantum theory and general relativity have modified this foundation in depth. The two theories have obtained solid success and vast experimental corroboration, and can be now considered as established knowledge. Each of the two theories modifies the conceptual foundation of classical physics in a (more or less) internally consistent manner, but we do not have a novel conceptual foundation capable of supporting both theories. This is why we do not yet have a theory capable of predicting what happens in the physical regime in which both theories are relevant, the regime of Planck scale phenomena, $10^{-33}$ cm.

General relativity has taught us not only that space and time share the property of being dynamical with the rest of the physical entities, but also (more crucially) that space–time location is relational only. Quantum mechanics has taught us that any dynamical entity is subject to Heisenberg’s uncertainty at small scale. Thus, we need a relational notion of a quan-
tum space–time, in order to understand Planck scale physics. Thus, for a relativist, the problem of quantum gravity is the problem of bringing a vast conceptual revolution, started with quantum mechanics and with general relativity, to a conclusion and to a new synthesis (see Rovelli (1997) Smolin (1997b)). In this synthesis, the notions of space and time need to be deeply reshaped, in order to keep into account what we have learned with both our present ‘fundamental’ theories.

Unlike perturbative or non–perturbative string theory, loop quantum gravity is formulated without a background space–time, and is thus a genuine attempt to grasp what is quantum space–time at the fundamental level. Accordingly, the notion of space–time that emerges from the theory is profoundly different from the one on which conventional quantum field theory or string theory are based.

According to Rovelli, the main merit of string theory is that it provides a superbly elegant unification of known fundamental physics, and that it has a well defined perturbation expansion, finite order by order. Its main incompleteness is that its non–perturbative regime is poorly understood, and that we do not have a background–independent formulation of the string theory. In a sense, we do not really know what is the theory we are talking about. Because of this poor understanding of the non perturbative regime of the theory, Planck scale physics and genuine quantum gravitational phenomena are not easily controlled: except for a few computations, there is not much Planck scale physics derived from string theory so far. There are, however, two sets of remarkable physical results. The first is given by some very high energy scattering amplitudes that have been computed. An intriguing aspect of these results is that they indirectly suggest that geometry below the Planck scale cannot be probed –and thus in a sense does not exist– in string theory. The second physical achievement of string theory (which followed the D–branes revolution) is the derivation of the Bekenstein–Hawking black hole entropy formula for certain kinds of black holes.

On the other hand, the main merit of loop quantum gravity is that it provides a well–defined and mathematically rigorous formulation of a background–independent non–perturbative generally covariant quantum field theory. The theory provides a physical picture and quantitative predictions on the world at the Planck scale. The main incompleteness of the theory regards the dynamics, formulated in several variants. The theory has lead to two main sets of physical results. The first is the derivation of the (Planck scale) eigenvalues of geometrical quantities such as areas and
volumes. The second is the derivation of black hole entropy for ‘normal’ black holes (but only up to the precise numerical factor).

The main physical hypotheses on which loop quantum gravity relies are only general relativity and quantum mechanics. In other words, loop quantum gravity is a rather conservative ‘quantization’ of general relativity, with its traditional matter couplings. In this sense, it is very different from string theory, which is based on a strong physical hypothesis with no direct experimental support ‘that the world is made by strings’.

Finally, strings and loop gravity, may not necessarily be competing theories: there might be a sort of complementarity, at least methodological, between the two. This is due to the fact that the open problems of string theory regard its background–independent formulation, and loop quantum gravity is precisely a set of techniques for dealing non–perturbatively with background independent theories. Perhaps the two approaches might even, to some extent, converge. Undoubtedly, there are similarities between the two theories: first of all the obvious fact that both theories start with the idea that the relevant excitations at the Planck scale are one dimensional objects – call them loops or strings. Smolin (1997a) also explored the possible relations between string theory and loop quantum gravity.

Loop quantum gravity is a quantum field theory on a differentiable 4–manifold. We have learned with general relativity that the space–time metric and the gravitational field are the same physical entity. Thus, a quantum theory of the gravitational field is a quantum theory of the space–time metric as well. It follows that quantum gravity cannot be formulated as a quantum field theory over a metric manifold, because there is no (classical) metric manifold whatsoever in a regime in which gravity (and therefore the metric) is a quantum variable Rovelli (1998).

One could conventionally split the space–time metric into two terms: one to be consider as a background, which gives a metric structure to space–time; the other to be treated as a fluctuating quantum field. This, indeed, is the procedure on which old perturbative quantum gravity, perturbative strings, as well as current non-perturbative string theories (M–theory), are based. In following this path, one assumes, for instance, that the causal structure of space–time is determined by the underlying background metric alone, and not by the full metric. Contrary to this, in loop quantum gravity we assume that the identification between the gravitational field and the metric–causal structure of space–time holds, and must be taken into account, in the quantum regime as well. Thus, no split of the metric is made, and there is no background metric on space–time.
We can still describe space–time as a (differentiable) manifold (a space without metric structure), over which quantum fields are defined. A classical metric structure will then be defined by expectation values of the gravitational field operator. Thus, the problem of quantum gravity is the problem of understanding what is a quantum field theory on a manifold, as opposed to quantum field theory on a metric space. This is what gives quantum gravity its distinctive flavor, so different than ordinary quantum field theory. In all versions of ordinary quantum field theory, the metric of space–time plays an essential role in the construction of the basic theoretical tools (creation and annihilation operators, canonical commutation relations, gaussian measures, propagators); these tools cannot be used in quantum field over a manifold.

Technically, the difficulty due to the absence of a background metric is circumvented in loop quantum gravity by defining the quantum theory as a representation of a Poisson algebra of classical observables, which can be defined without using a background metric. The idea that the quantum algebra at the basis of quantum gravity is not the canonical commutation relation algebra, but the Poisson algebra of a different set of observables has long been advocated by [Isham (1984)], whose ideas have been very influential in the birth of loop quantum gravity. The algebra on which loop gravity is the loop algebra [Rovelli and Smolin (1990)].

In choosing the loop algebra as the basis for the quantization, we are essentially assuming that Wilson loop operators are well defined in the Hilbert space of the theory. In other words, that certain states concentrated on one dimensional structures (loops and graphs) have finite norm. This is a subtle non trivial assumptions entering the theory. It is the key assumption that characterizes loop gravity. If the approach turned out to be wrong, it will likely be because this assumption is wrong. The Hilbert space resulting from adopting this assumption is not a Fock space. Physically, the assumption corresponds to the idea that quantum states can be decomposed on a basis of Faraday lines–excitations (as Minkowski QFT states can be decomposed on a particle basis).

Furthermore, this is an assumption that fails in conventional quantum field theory, because in that context well defined operators and finite norm states need to be smeared in at least three dimensions, and 1D objects are too singular. The fact that at the basis of loop gravity there is a mathematical assumption that fails for conventional Yang–Mills quantum field theory is probably at the origin of some of the resistance that loop quantum gravity encounters among some high energy theorists. What distinguishes gravity
from Yang–Mills (YM) theories, however, and makes this assumption viable in gravity even if it fails for YM theory is *diffeomorphism invariance*. The loop states are singular states that span a ‘huge’ non–separable state space. Non–perturbative diffeomorphism invariance plays two roles. First, it wipes away the infinite redundancy. Second, it ‘smears’ a loop state into a knot state, so that the physical states are not really concentrated in one dimension, but are, in a sense, smeared all over the entire manifold by the non–perturbative diffeomorphisms [Rovelli (1998)].

Conventional field theories are *not invariant* under a diffeomorphism acting on the dynamical fields. Every field theory, suitably formulated, is trivially invariant under a diffeomorphism acting on *everything*. General relativity, on the contrary is invariant under such transformations. More precisely, every general relativistic theory has this property. Thus, diffeomorphism invariance is not a feature of just the gravitational field: it is a feature of physics, once the existence of relativistic gravity is taken into account. Thus, one can say that the gravitational field is not particularly ‘special’ in this regard, but that diffeomorphism invariance is a property of the physical world that can be disregarded only in the approximation in which the dynamics of gravity is neglected.

Now, *diffeomorphism invariance* is the technical implementation of a physical idea, due to Einstein. The idea is a deep modification of the pre–general–relativistic (pre–GR) notions of space and time. In pre–GR physics, we assume that physical objects can be localized in space and time with respect to a fixed non–dynamical background structure. Operationally, this background space–time can be defined by means of physical reference–system objects, but these objects are considered as dynamically decoupled from the physical system that one studies. This conceptual structure fails in a relativistic gravitational regime. In general relativistic physics, the physical objects are localized in space and time only with respect to each other. Therefore if we ‘displace’ all dynamical objects in space–time at once, we are not generating a different state, but an equivalent mathematical description of the same physical state. Hence, diffeomorphism invariance.

Accordingly, a physical state in GR is not ‘located’ somewhere. Pictorially, GR is not physics over a stage, it is the dynamical theory of (or including) the stage itself. Loop quantum gravity is an attempt to implement this subtle relational notion of space–time localization in quantum field theory. In particular, the basic quantum field theoretical excitations cannot be localized somewhere as, say, photons are. They are quantum excitations of the ‘stage’ itself, not excitations over a stage. Intuitively,
one can understand from this discussion how knot theory plays a role in the theory. First, we define quantum states that correspond to loop–like excitations of the gravitational field, but then, when factoring away diffeomorphism invariance, the location of the loop becomes irrelevant. The only remaining information contained in the loop is then its knotting (a knot is a loop up to its location). Thus, diffeomorphism invariant physical states are labelled by knots. A knot represent an elementary quantum excitation of space. It is not here or there, since it is the space with respect to which here and there can be defined. A knot state is an elementary quantum of space. In this manner, loop quantum gravity ties the new notion of space and time introduced by general relativity with quantum mechanics.

4.2.2 Formalism of Loop Quantum Gravity

The starting point is classical general relativity formulated in terms of the Ashtekar phase–space formalism (see [Ashtekar (1991)]). Recall that classical general relativity can be formulated in the phase–space form as follows.

We fix a 3D manifold \( M \) (compact and without boundaries) and consider a smooth real \( SU(2) \)–connection \( A^i_a(x) \) and a vector density \( \tilde{E}^a_i(x) \) (transforming in the vector representation of \( SU(2) \)) on \( M \). We use \( a, b, \ldots = 1, 2, 3 \) for spatial indices and \( i, j, \ldots = 1, 2, 3 \) for internal indices. The internal indices can be viewed as labelling a basis in the Lie algebra of \( SU(2) \) or the three axis of a local triad. We indicate coordinates on \( M \) with \( x \). The relation between these fields and conventional metric gravitational variables is as follows: \( \tilde{E}^a_i(x) \) is the (densitized) inverse triad, related to the 3D metric \( g_{ab}(x) \) of constant–time surfaces by

\[
\tilde{E}^a_i(x) \tilde{E}^b_j(x) = g_{ij}(x),
\]

where \( g \) is the determinant of \( g_{ab} \); and

\[
A^i_a(x) = \Gamma^i_a(x) + \gamma k^i_a(x);
\]

\( \Gamma^i_a(x) \) is the spin connection associated to the triad, (defined by \( \partial_a e^i_b = \Gamma^i_a e^a_b \), where \( e^i_a \) is the triad).

\( k^i_a(x) \) is the extrinsic curvature of the constant time three surface.

In (4.15), \( \gamma \) is a constant, denoted the Immirzi parameter, that can be chosen arbitrarily (it will enter the Hamiltonian constraint). Different choices for \( \gamma \) yield different versions of the formalism, all equivalent in the classical domain. If we choose \( \gamma \) to be equal to the imaginary unit, \( \gamma = \sqrt{-1} \), then \( A \) is the standard Ashtekar connection, which can be shown...
to be the projection of the self-dual part of the 4D spin connection on the constant time surface. If we choose $\gamma = 1$, we get the real Barbero connection. The Hamiltonian constraint of Lorentzian general relativity has a particularly simple form in the $\gamma = \sqrt{-1}$ formalism, while the Hamiltonian constraint of Euclidean general relativity has a simple form when expressed in terms of the $\gamma = 1$ real connection. Other choices of $\gamma$ are viable as well. In particular, it has been argued that the quantum theory based on different choices of $\gamma$ are genuinely physical inequivalent, because they yield ‘geometrical quanta’ of different magnitude [Rovelli (1998)]. Apparently, there is a unique choice of $\gamma$ yielding the correct $1/4$ coefficient in the Bekenstein–Hawking formula.

The spinorial version of the Ashtekar variables is given in terms of the Pauli matrices $\sigma_i, i = 1, 2, 3$, or the $\mathfrak{su}(2)$ generators $\tau_i = -\frac{i}{2} \sigma_i$, by

$$\tilde{E}^a(x) = -i \tilde{E}^a_i(x) \sigma_i = 2 \tilde{E}^a_i(x) \tau_i, \quad (4.16)$$
$$A_a(x) = -\frac{i}{2} A^i_a(x) \sigma_i = A^i_a(x) \tau_i. \quad (4.17)$$

Thus, $A_a(x)$ and $\tilde{E}^a(x)$ are $2 \times 2$ anti-Hermitian complex matrices.

The theory is invariant under local $SU(2)$ gauge transformations, three-dimensional diffeomorphisms of the manifold on which the fields are defined, as well as under (coordinate) time translations generated by the Hamiltonian constraint. The full dynamical content of general relativity is captured by the three constraints that generate these gauge invariances (see [Ashtekar (1991)]).

### 4.2.3 Loop Algebra

Certain classical quantities play a very important role in the quantum theory. These are: the trace of the holonomy of the connection, which is labelled by loops on the three manifold; and the higher order loop variables, obtained by inserting the $E$ field (in $n$ distinct points, or ‘hands’) into the holonomy trace. More precisely, given a loop $\alpha$ in $M$ and the points $s_1, s_2, \ldots, s_n \in \alpha$ we define:

$$T[\alpha] = -\text{Tr}[U_\alpha], \quad (4.18)$$
$$T^a[\alpha](s) = -\text{Tr}[U_\alpha(s, s) \tilde{E}^a(s)] \quad (4.19)$$
and, in general
\[
T^a_{s_1 s_2} \alpha = - \text{Tr}[U_\alpha(s_1, s_2) \tilde{E}^a_{s_2} U_\alpha(s_2, s_1) \tilde{E}^\alpha(s_1)],
\]
(4.20)
\[
T^a_1 \cdots a_N \alpha(s_1 \ldots s_N) = - \text{Tr}[U_\alpha(s_1, s_N) \tilde{E}^a_{s_N} U_\alpha(s_N, s_{N-1}) \cdots \tilde{E}^\alpha(s_1)]
\]
where \( U_\alpha(s_1, s_2) \sim \mathcal{P} \exp \{ \int_{s_2}^{s_1} A^a(\alpha(s)) \text{d}s \} \) is the parallel propagator of \( A_a \) along \( \alpha \), defined by
\[
\frac{d}{ds} U_\alpha(1, s) = \frac{d}{ds} A_a(\alpha(s)) U_\alpha(1, s).
\]
(4.21)
These are the loop observables, previously introduced in YM theories.

The loop observables coordinate the phase space and have a closed Poisson algebra, denoted the loop algebra. This algebra has a remarkable geometrical flavor. For instance, the Poisson bracket between \( T^a_\alpha \) and \( T^a_\beta(s) \) is non vanishing only if \( \beta(s) \) lies over \( \alpha \); if it does, the result is proportional to the holonomy of the Wilson loops obtained by joining \( \alpha \) and \( \beta \) at their intersection (by rerouting the 4 legs at the intersection). More precisely
\[
\{ T^a_\alpha, T^a_\beta(s) \} = \Delta^a_\alpha(\beta) \cdot [T^a_\alpha \# \beta - T^a_\alpha \# \beta^{-1}].
\]
(4.22)
Here
\[
\Delta^a_\alpha(\beta) = \int ds \frac{d\alpha(s)}{ds} \delta^3(\alpha(s), \beta)
\]
(4.23)
is a vector distribution with support on \( \alpha \) and \( \alpha \# \beta \) is the loop obtained starting at the intersection between \( \alpha \) and \( \beta \), and following first \( \alpha \) and then \( \beta \), \( \beta^{-1} \) is \( \beta \) with reversed orientation.

A (non–SU(2) gauge invariant) quantity that plays a role in certain aspects of the theory, particularly in the regularization of certain operators, is obtained by integrating the \( E \) field over a two dimensional surface \( S \)
\[
E[S, f] = \int_S dS_a E^a_i f^i,
\]
(4.24)
where \( f \) is a function on the surface \( S \), taking values in the Lie algebra of \( SU(2) \). In alternative to the full loop observables \( (4.18), (4.19), (4.20) \), one also can take the holonomies and \( E[S, f] \) as elementary variables.

4.2.4 Loop Quantum Gravity

The kinematic of a quantum theory is defined by an algebra of ‘elementary’ operators (such as \( x \) and \( i\hbar d/dx \), or creation and annihilation operators)
on a Hilbert space $\mathcal{H}$. The physical interpretation of the theory is based on the connection between these operators and classical variables, and on the interpretation of $\mathcal{H}$ as the space of the quantum states. The dynamics is governed by a Hamiltonian, or, as in general relativity, by a set of quantum constraints, constructed in terms of the elementary operators. To assure that the quantum Heisenberg equations have the correct classical limit, the algebra of the elementary operator has to be isomorphic to the Poisson algebra of the elementary observables. This yields the 

heuristic quantization rule: `promote Poisson brackets to commutators’. In other words, define the quantum theory as a linear representation of the Poisson algebra formed by the elementary observables. The kinematics of the quantum theory is defined by a unitary representation of the loop algebra.

We can start à la Schrödinger, by expressing quantum states by means of the amplitude of the connection, namely by means of functionals $\Psi(A)$ of the (smooth) connection. These functionals form a linear space, which we promote to a Hilbert space by defining a inner product. To define the inner product, we choose a particular set of states, which we denote ‘cylindrical states’ and begin by defining the scalar product between these.

Pick a graph $\Gamma$, say with $n$ links, denoted $\gamma_1 \ldots \gamma_n$, immersed in the manifold $M$. For technical reasons, we require the links to be analytic. Let $U_i(A) = U_{\gamma_i}$, $i = 1, \ldots, n$ be the parallel transport operator of the connection $A$ along $\gamma_i$. $U_i(A)$ is an element of $SU(2)$. Pick a function $f(g_1 \ldots g_n)$ on $[SU(2)]^n$. The graph $\Gamma$ and the function $f$ determine a functional of the connection as follows

$$\psi_{\Gamma,f}(A) = f(U_1(A), \ldots, U_n(A)), \quad (4.25)$$

(these states are called cylindrical states because they were previously introduced as cylindrical functions for the definition of a cylindrical measure). Notice that we can always ‘enlarge the graph’, in the sense that if $\Gamma$ is a subgraph of $\Gamma'$ we can write

$$\psi_{\Gamma,f}(A) = \psi_{\Gamma',f'}(A), \quad (4.26)$$

by simply choosing $f'$ independent from the $U_i$’s of the links which are in $\Gamma'$ but not in $\Gamma$. Thus, given any two cylindrical functions, we can always view them as having the same graph (formed by the union of the two graphs). Given this observation, we define the scalar product between any...
two cylindrical functions, by

\[ (\psi_{\Gamma,f}, \psi_{\Gamma,h}) = \int_{SU(2)^n} dg_1 \cdots dg_n f(g_1 \cdots g_n) h(g_1 \cdots g_n) \]  

(4.27)

where \( dg \) is the Haar measure on \( SU(2) \). This scalar product extends by linearity to finite linear combinations of cylindrical functions. It is not difficult to show that (4.27) defines a well defined scalar product on the space of these linear combinations. Completing the space of these linear combinations in the Hilbert norm, we get a Hilbert space \( \mathcal{H} \). This is the (unconstrained) quantum state space of loop gravity. \( \mathcal{H} \) carries a natural unitary representation of the diffeomorphism group and of the group of the local \( SU(2) \) transformations, obtained transforming the argument of the functionals. An important property of the scalar product (4.27) is that it is invariant under both these transformations.

\( \mathcal{H} \) is non-separable. At first sight, this may seem as a serious obstacle for its physical interpretation. But we will see below that after factoring away diffeomorphism invariance we may get a separable Hilbert space. Also, standard spectral theory holds on \( \mathcal{H} \), and it turns out that using spin networks (discussed below) one can express \( \mathcal{H} \) as a direct sum over finite dimensional subspaces which have the structure of Hilbert spaces of spin systems; this makes practical calculations very manageable.

Finally, we will use a Dirac notation and write

\[ \Psi(A) = \langle A | \Psi \rangle, \]  

(4.28)

in the same manner in which one may write \( \psi(x) = \langle x | \Psi \rangle \) in ordinary quantum mechanics. As in that case, however, we should remember that \( |A \rangle \) is not a normalizable state.

### 4.2.5 Loop States and Spin Network States

A subspace \( \mathcal{H}_0 \) of \( \mathcal{H} \) is formed by states invariant under \( SU(2) \) gauge transformations. We now define an orthonormal basis in \( \mathcal{H}_0 \). This basis represents a very important tool for using the theory. It was in-
introduced in [Rovelli and Smolin (1995)] and developed in [Baez (1996a); Baez (1996b)], it is denoted spin network basis.

First, given a loop $\alpha$ in $M$, there is a normalized state $\psi_\alpha(A)$ in $\mathcal{H}$, which is obtained by taking $\Gamma = \alpha$ and $f(g) = -\text{Tr}(g)$. Namely

$$\psi_\alpha(A) = -\text{Tr}(U_\alpha(A)).$$  \hspace{1cm} (4.29)

We introduce a Dirac notation for the abstract states, and denote this state as $|\alpha\rangle$. These states are called loop states. Using Dirac notation, we can write

$$\psi_\alpha(A) = \langle A|\alpha\rangle.$$  \hspace{1cm} (4.30)

It is easy to show that loop states are normalizable. Products of loop states are normalizable as well. Following tradition, we denote with $\alpha$ also a multi–loop, namely a collection of (possibly overlapping) loops $\{\alpha_1, \ldots, \alpha_n\}$, and we call

$$\psi_\alpha(A) = \psi_{\alpha_1}(A) \times \ldots \times \psi_{\alpha_n}(A)$$  \hspace{1cm} (4.31)

- a multi–loop state. Multi–loop states represented the main tool for loop quantum gravity before the discovery of the spin network basis. Linear combinations of multi–loop states over–span $\mathcal{H}$, and therefore a generic state $\psi(A)$ is fully characterized by its projections on the multi–loop states, namely by

$$\psi(\alpha) = \langle \psi_\alpha, \psi \rangle.$$  \hspace{1cm} (4.32)

The ‘old’ loop representation was based on representing quantum states in this manner, namely by means of the functionals $\psi(\alpha)$ over loop space defined in (4.32).

Next, consider a graph $\Gamma$. A ‘coloring’ of $\Gamma$ is given by the following.

1. Associate an irreducible representation of $SU(2)$ to each link of $\Gamma$. Equivalently, we may associate to each link $\gamma_i$ a half integer number $s_i$, the spin of the irreducible, or, equivalently, an integer number $p_i$, the ‘color’ $p_i = 2s_i$.

2. Associate an invariant tensor $v$ in the tensor product of the representations $s_1 \ldots s_n$, to each node of $\Gamma$ in which links with spins $s_1 \ldots s_n$ meet. An invariant tensor is an object with $n$ indices in the representations $s_1 \ldots s_n$ that transform covariantly. If $n = 3$, there is only one invariant tensor (up to a multiplicative factor), given by the Clebsh–Gordon coefficient. An invariant tensor is also called an intertwining tensor.
All invariant tensors are given by the standard Clebsch–Gordon theory. More precisely, for fixed $s_1 \ldots s_n$, the invariant tensors form a finite dimensional linear space. Pick a basis $v_j$ in this space, and associate one of these basis elements to the node. Notice that invariant tensors exist only if the tensor product of the representations $s_1 \ldots s_n$ contains the trivial representation. This yields a condition on the coloring of the links. For $n = 3$, this is given by the well known Clebsh–Gordan condition: each color is not larger than the sum of the other two, and the sum of the three colors is even.

We indicate a colored graph by $\{\Gamma, \vec{s}, \vec{v}\}$, or simply $S = \{\Gamma, \vec{s}, \vec{v}\}$, and denote it a ‘spin network’. (It was R. Penrose who first had the intuition that this mathematics could be relevant for describing the quantum properties of the geometry, and who gave the first version of spin network theory [Penrose (1971a); Penrose (1971b)].)

Given a spin network $S$, we can construct a state $\Psi_S(A)$ as follows. We take the propagator of the connection along each link of the graph, in the representation associated to that link, and then, at each node, we contract the matrices of the representation with the invariant tensor. We get a state $\Psi_S(A)$, which we also write as

$$\psi_S(A) = \langle A|S \rangle. \quad (4.33)$$

One can then show the following.

- The spin network states are normalizable. The normalization factor is computed in [DePietri and Rovelli (1996)].
- They are $SU(2)$ gauge invariant.
- Each spin network state can be decomposed into a finite linear combination of products of loop states.
- The (normalized) spin network states form an orthonormal basis for the gauge $SU(2)$ invariant states in $\mathcal{H}$ (choosing the basis of invariant tensors appropriately).
- The scalar product between two spin network states can be easily computed graphically and algebraically.

The spin network states provide a very convenient basis for the quantum theory.

The spin network states defined above are $SU(2)$ gauge invariant. There exists also an extension of the spin network basis to the full Hilbert space.
4.2.6 **Diagrammatic Representation of the States**

A diagrammatic representation for the states in $\mathcal{H}$ is very useful in concrete calculations. First, associate to a loop state $|\alpha\rangle$ a diagram in $M$, formed by the loop $\alpha$ itself. Next, notice that we can *multiply* two loop states, obtaining a normalizable state. We represent the product of $n$ loop states by the diagram formed by the set of the $n$ (possibly overlapping) corresponding loops (we denote this set ‘multi–loop’). Thus, linear combinations of multi–loops diagrams represent states in $\mathcal{H}$. Representing states as linear combinations of multi–loops diagrams makes computation in $\mathcal{H}$ easy.

Now, the spin network state defined by the graph with no nodes $\alpha$, with color 1, is clearly, by definition, the loop state $|\alpha\rangle$, and we represent it by the diagram $\alpha$. The spin network state $|\alpha, n\rangle$ determined by the graph without nodes $\alpha$, with color $n$ can be obtained as follows. Draw $n$ parallel lines along the loop $\alpha$; cut all lines at an arbitrary point of $\alpha$, and consider the $n!$ diagrams obtained by joining the legs after a permutation. The linear combination of these $n!$ diagrams, taken with alternate signs (namely with the sign determined by the parity of the permutation) is precisely the state $|\alpha, n\rangle$. The reason of this key result can be found in the fact that an irreducible representation of $SU(2)$ can be obtained as the totally symmetric tensor product of the fundamental representation with itself [DePietri and Rovelli (1996)].

Next, consider a graph $\Gamma$ with nodes. Draw $n_i$ parallel lines along each link $\gamma_i$. Join pairwise the end points of these lines at each node (in an arbitrary manner), in such a way that each line is joined with a line from a different link. In this manner, one get a multi–loop diagram. Now antisymmetrize the $n_i$ parallel lines along each link, obtaining a linear combination of diagrams representing a state in $\mathcal{H}$. One can show that this state is a spin network state, where $n_i$ is the color of the links and the color of the nodes is determined by the pairwise joining of the legs chosen [DePietri and Rovelli (1996)]. Again, simple $SU(2)$ representation theory is behind this result.

More in detail, if a node is trivalent (has 3 adjacent links), then we can join legs pairwise only if the total number of the legs is even, and if the number of the legs in each link is smaller or equal than the sum of the number of the other two. This can be immediately recognized as the **Clebsch–Gordan condition**. If these conditions are satisfied, there is only a single way of joining legs. This corresponds to the fact that there is only one invariant tensor in the product of three irreducible of $SU(2)$. Higher
valence nodes can be decomposed into trivalent ‘virtual’ nodes, joined by ‘virtual’ links. Orthogonal independent invariant tensors are obtained by varying over all allowed colorings of these virtual links (compatible with the Clebsch–Gordan conditions at the virtual nodes). Different decompositions of the node give different orthogonal bases. Thus the total (links and nodes) coloring of a spin network can be represented by means of the coloring of the real and the virtual links (see Figure 4.2).

Vice versa, multi-loop states can be decomposed into spin network states by simply symmetrizing along (real and virtual) nodes. This can be done particularly easily diagrammatically, as illustrated by the graphical formulae in [Rovelli and Smolin (1995); DePietri and Rovelli (1996)]. These are standard formulae. In fact, it is well known that the tensor algebra of the $SU(2)$ irreducible representations admits a completely graphical notation. This graphical notation has been widely used for instance in nuclear and atomic physics. The application of this diagrammatic calculus to quantum gravity is described in detail in [DePietri and Rovelli (1996)].

### 4.2.7 Quantum Operators

Now, we define the quantum operators, corresponding to the $T$-variables, as linear operators on $\mathcal{H}$. These form a representation of the loop variables Poisson algebra. The operator $T[\alpha]$ acts diagonally

$$ T[\alpha] \Psi(A) = -\text{Tr} U_\alpha(A) \Psi(A), $$

(recall that products of loop states and spin-network states are normalizable states). In diagrammatic notation, the operator simply adds a loop to
a (linear combination of) multi–loops

\[ T[\alpha]|\Psi\rangle = |\alpha\rangle|\Psi\rangle. \]

Higher order loop operators are expressed in terms of the elementary ‘grasp’ operation. Consider first the operator \( T^\alpha(s)[\alpha] \), with one hand in the point \( \alpha(s) \). The operator annihilates all loop states that do not cross the point \( \alpha(s) \). Acting on a loop state \( |\beta\rangle \), it gives

\[ T^\alpha(s)[\alpha]|\beta\rangle = l_0^2 \Delta^\alpha[\beta,\alpha(s)][|\alpha\#\beta\rangle - |\alpha\#\beta^{-1}\rangle], \]

where we have introduced the elementary length \( l_0 \) by

\[ l_0^2 = \hbar G = \frac{16\pi \hbar G_{\text{Newton}}}{c^3} = 16\pi l_{\text{Planck}}^2 \]

and \( \Delta^\alpha \) and \# were defined above. This action extends by linearity, continuity and by the Leibniz rule to products and linear combinations of loop states, and to the full \( \mathcal{H} \). In particular, it is not difficult to compute its action on a spin network state [DePietri and Rovelli (1996)]. Higher order loop operators act similarly. It is easy to verify that these operators provide a representation of the classical Poisson loop algebra.

All the operators in the theory are then constructed in terms of these basics loop operators, in the same way in which in conventional QFT one constructs all operators, including the Hamiltonian, in terms of creation and annihilation operators. The construction of the composite operators requires the development of regularization techniques that can be used in the absence of a background metric.

### 4.2.8 Loop v.s. Connection Representation

Imagine we want to quantize the one dimensional harmonic oscillator. We can consider the Hilbert space of square integrable functions \( \psi(x) \) on the real line, and express the momentum and the Hamiltonian as differential operators. Denote the eigenstates of the Hamiltonian as \( \psi_n(x) = \langle x|n \rangle \). It is well known that the theory can be expressed entirely in algebraic form in terms of the states \( |n\rangle \). In doing so, all elementary operators are algebraic:

\[ \hat{x}|n\rangle = \frac{1}{\sqrt{2}}(|n-1\rangle + (n+1)|n+1\rangle), \quad \hat{p}|n\rangle = \frac{\sqrt{2}}{2^n}((n-1) - (n+1)|n+1\rangle). \]

Similarly, in quantum gravity we can directly construct the quantum theory in the spin–network (or loop) basis, without ever mentioning functionals of the connections. This representation of the theory is denoted the loop representation.
A section of the first paper on loop quantum gravity by Rovelli and Smolin (1990) was devoted to a detailed study of ‘transformation theory’ (in the sense of Dirac) on the state space of quantum gravity, and in particular on the relations between the loop states

$$\psi(\alpha) = \langle \alpha | \psi \rangle$$

(4.36)

and the states $\psi(A)$ giving the amplitude for a connection field configuration $A$, and defined by

$$\psi(A) = \langle A | \psi \rangle.$$ (4.37)

Here $|A\rangle$ are ‘eigenstates of the connection operator’, or, more precisely (since the operator corresponding to the connection is ill defined in the theory) the generalized states that satisfy

$$T[\alpha] |A\rangle = -\text{Tr}[Pe^{iA}] |A\rangle.$$ (4.38)

However, at the time of Rovelli and Smolin (1990) the lack of a scalar product made transformation theory quite involved.

On the other hand, the introduction of the scalar product (4.27) gives a rigorous meaning to the loop transform. In fact, we can write, for every spin network $S$, and every state $\psi(A)$

$$\psi(S) = \langle S | \psi \rangle = (\psi_S, \psi).$$ (4.39)

This equation defines a unitary mapping between the two presentations of $\mathcal{H}$: the ‘loop representation’, in which one works in terms of the basis $|S\rangle$; and the ‘connection representation’, in which one uses wave functionals $\psi(A)$.

The development of the connection representation followed a winding path through $C^*$-algebraic and measure theoretical methods. The work of DePietri (1997) has proven the unitary equivalence of the two formalisms.

### 4.3 Quantum Cosmology

In this section we review the pinnacle of modern physical science, the cosmological dynamics, mainly following the strongest player in this exciting research field, Stephen Hawking.
4.3.1 Hawking’s Cosmology in ‘Plain English’

According to S. Hawking and his new collaborator T. Hertog, there is no one history of the universe. There is no immutable past, no 13.7 billion years of evolution for cosmologists to retrace. Instead, there are many possible histories, and the universe has lived them all [Gefter (2006)].

All this started in Hawking’s early work with mathematical physicist R. Penrose in 1970s. They proved their Singularity Theorems, effectively showing that our expanding universe must have emerged from a black-hole like singularity, a point of infinite curvature, a place where gravity becomes so strong that space and time are curved beyond recognition, where general relativity (our best description of how space, time and matter interact) – no longer applied. Now, instead of gravitation theory, Hawking and Hertog suggest that the universe was so small at this time that quantum effects must have been important. Hertog claims, “The real lesson of these Singularity theorems is that the origin of the universe is a quantum event.”

In 1983, Hawking and J. Hartle took the picture of the famous quantum double-slit experiment and applied it to the evolution of the whole universe, using Feynman’s sum-over-histories interpretation of quantum physics, i.e., path integral methodology. Recall that R. Feynman suggested that the way to interpret quantum phenomena, such as the double-slit experiment, was to assume that when a particle travels from point A to point B, it doesn’t simply take one path; it rather takes every possible path simultaneously; e.g., the photon travels through both slits at the same time and interferes with itself. In this scheme, when a photon travels from a lam to our eye it moves in a straight line, but it also dances about in twists and swirls. The obvious question, then, is why do we ever see only one path, straight and simple? Feynman’s answer was, because all the other paths cancel each other out. In the sum-over-histories interpretation, each path, or history, can be mapped out as a wave. Each wave has a different phase (effectively a starting time), and all the waves added together create an ‘interference pattern’, building upon one another where their phases align and cancelling each other out where their phases are mismatched. The sum of all the waves is one single wave, the so-called wave $\psi$–function, which describes the path we observe.

Applied to the universe, this idea has an obvious implication [Gefter (2006)]. Just as a particle travelling from point A to point B takes every possible path in between, so too must the history of the universe. Hertog says, “The universe doesn’t have a single history, but every possible history,
each with its own probability.” This is Hawking’s famous wave function of the universe.

The Wave Function of the Universe

This approach starts with the idea of not just our universe but all possible universes; then calculates the relative likelihoods of each of them, much like calculating the wave \( \psi \)–function of a particle. Recall that the wave function \( \psi = \psi(t) \) of a particle is essentially the function that determines its most likely location at any given time. Wave–functions are largest where that particle is observed to be, but also extend throughout the known universe in accordance with the sum–over–histories method. Everything has a wave \( \psi \)–function: elementary particles possess wave–functions and make up all other matter in the universe, so this is a logical conclusion. Large and common objects such as rubber balls have wave–functions; for example, the wave–function of a ball sitting on a flat surface is largest where it is observed – say, a table – but also extends everywhere around us, or on the moon, or even in another galaxy. However, the likelihood of the ball suddenly appearing in any of these locations is infinitesimal. The likelihood of such changes in location depends on Planck’s constant \( \hbar \).

Hawking and Hartle have proposed to calculate the wave–function of the universe using the sum–over–histories method, which begins with the assumption that the universe has all possible histories. Moreover, they would calculate this sum in imaginary time, not ordinary time. This is because imaginary time travels at right angles to ordinary time and ‘meets’ with the three spatial dimensions to create a smooth surface similar to the surface of the earth. This eliminates the singularities (points of infinite curvature) present in ordinary time, allowing the history of the universe to be reliably calculated. Also unlike ordinary time, imaginary time has no beginning or end, so progression through it is determined entirely by physical laws.

For Hawking and Hartle’s calculation, we must begin with a wave–function describing all possible universes – an infinite number of them. The wave–function is large near our own universe and infinitesimal near others in which life is impossible or the known laws of physics do not apply. Because of the wave–function’s concentration in our own universe, it is the most likely of them all, but there is a chance (albeit vanishingly small) that an object from this universe would suddenly make a quantum leap into another one. Proving this conjecture mathematically is one of the
primary goals of quantum cosmology, which applies quantum theory to the large structures of cosmology. The Hawking–Hartle theory also postulates the existence of wormholes connecting the different universes. According to them, the multitude of universes should be connected by wormholes, although these wormholes are not an efficient or readily available means of transportation. Some of these universes are very rich, and others are quite barren. Similarly, some are connected with many others, while others are isolated.

The wave–function of Hawking and Hartle raises two major controversies long debated among scientists. The first of these is an apparent return to the so–called anthropic principle (of S. Weinberg), which basically says that the universe is the way it is because we wouldn’t be here if it was any other way. It has two basic forms: the ‘weak’ and ‘strong’ versions. The weak version states that the existence of intelligent life is experimental evidence to help us understand the universe’s seemingly random physical constants. The strong version, much more controversial, states that these apparently random constants are not random but were instead chosen by some Supreme Being to make life in our universe possible. Hawking and Hartle’s idea appeals to the weak anthropic principle, but apparently states that Supreme Being is not necessary to explain the existence of a universe uniquely suited to intelligent life. The second controversy relates to the famous Schrödinger’s cat paradox and the many–worlds theories. The Schrödinger’s cat paradox states that, by the Heisenberg uncertainty principle, a particle is in a sum of all possible states until it is observed, a process called reducing the wave–function reduction. Schrödinger postulated a paradox that arises by this theory: suppose a cat is placed in a box connected to a gun which is in turn connected to a Geiger counter measuring a uranium atom. If the unstable atom decays, the Geiger counter will register it, the gun will go off, and the cat will be killed. If it doesn’t decay, the process will not be initiated and the cat will live. Before observing the cat, quantum theory states that it is in a superimposed condition of both dead and live states. Most physicists either assume the wave–function is always being reduced by the cosmic observer, the Supreme Being, or simply ignore the problem. A third way of dealing with the paradox is the many–worlds theory, suggested by H. Everett, which states that the universe is constantly splitting off new offshoots, so that in one universe, the cat is dead, and in another it is alive. In the traditional many–worlds theory, contact between the worlds is mathematically impossible and therefore the idea cannot be tested. By the physics principle of Occam’s razor, which
states that the simpler of two competing explanations is generally correct, we should throw out the many-worlds theory because it is irrelevant to our universe and completely untestable. Hawking and Hartle’s idea revives the many-worlds theory with one important twist: communication between the worlds in the form of wormholes discussed above is possible in their formulation. Thus, their idea is both testable and directly relevant to our universe.

**Top-Down Cosmology with No-Boundary Proposal**

However, there is a twist: the history that we see depends on the experimental setup. Recall that in the double-slit experiment, if we use a photon detector to find which of the two slits the photon went through, it no longer creates an interference pattern, just a single spot on the film. In other words, the way we look at the photon changes the nature of its journey. The same thing happens in Hawking’s universe: our observations of the cosmos today are determining the outcome, that is the entire history of the universe. A measurement made in the present is deciding what happened 13.7 billion years ago; by looking out at the universe, we assign ourselves a particular, concrete history [Gefter (2006)].

Although this might look as a violation of the cause-and-effect laws, Hawking says that it is all to do with perspective. If we could stand outside the world, we would be able to to see the present affecting the past, as when an observer affects a photon’s path through the universe. From inside the universe, though, no observer sees causality violated. What we observe in the present, the ‘final’ state, is one entire, causally consistent history or another: from within any given history, cause and effect proceed in the usual manner. “Observations of final states determine histories of the universe. A worm’s-eye view from inside the universe would have the normal causality. Backwards causality is an angel’s view from outside the universe,” says Hawking.

So the idea is that to unravel the past, we must sum together all possible histories of the universe. Hawking and Hertog equate the cosmic histories with how the geometry (and topology) of the universe evolves in each possible case of going from the initial point A (the beginning of time) to the final point B (now). We can specify the state of the universe at the final point B by making certain observations of the world around us (e.g., the universe has three large spatial dimensions, its geometry is close to flat, it is expanding, etc.). However, what about the initial point A?
out the paths of a photon from a lamp to our eye is easy because we know the initial point A (the lamp) and the final point B (our eye). However, we know nothing about the universe at the beginning of time. And that is precisely what cosmology is supposed to tell us.

This is where the sum–over–histories interpretation comes into its own. Its mathematics contains an apparent oddity: the answers only come out right when the calculation is done in imaginary time. Hawking and Hartle’s original work on the wave function of the universe (i.e., quantum properties of the cosmos), suggested that imaginary time, which previously seemed like a mathematical curiosity in the sum–over–histories, held the answer to understanding the origin of the universe. Add up the histories of the universe in imaginary time, and time is transformed into space. The result is that, when the universe was small enough to be governed by quantum mechanics, it had four spatial dimensions and no dimension of time: where time would usually come to an end at a singularity, a new dimension of space appears and the singularity vanishes [Gefter (2006)].

In terms of the universe’s history, that means there is no initial point A; this is called the no–boundary proposal. “Like the surface of a sphere, our universe has no definable starting point.” This has led Hawking to define a new kind of cosmology. The traditional approach, which he calls bottom–up cosmology, tries to specify the initial state of the universe and work from there. This is doomed to fail, Hawking says, because we know nothing about the initial conditions. Instead, he suggests, we should use the no–boundary proposal to do top–down cosmology, where the only input into our models of the universe comes from what we observe now. The result of this process, he says, solves a long-standing problem of cosmology, called ‘fine-tuning’. For example, most cosmologists think that the universe went through an early burst of rapid expansion, the so–called inflation.\footnote{Recall that in physical cosmology, the inflation is the idea that the nascent universe passed through a phase of exponential expansion that was driven by a negative-pressure vacuum energy density. As a direct consequence of this expansion, all of the observable universe originated in a small causally–connected region. Inflation answers the classic conundrums of the Big–Bang cosmology: why does the universe appear flat, homogeneous and isotropic in accordance with the cosmological principle when one would expect, on the basis of the physics of the Big–Bang, a highly curved, inhomogeneous universe. Inflation also explains the origin of the large-scale structure of the cosmos. Quantum fluctuations in the microscopic inflationary region, magnified to cosmic size, become the seeds for the growth of structure in the universe (see galaxy formation and evolution and structure formation). While the detailed particle physics mechanism responsible for inflation is not known, the basic picture makes a number of predictions that have been confirmed by observational tests. Inflation is thus now considered part of the standard cosmology.} The problem here is that
standard inflationary models require a very improbable initial state, one that must have 'finely tuned' values that cause inflation to start, then stop in a certain way after a certain time: complicated prescription whose only justification is to produce a flat universe without any strange topology, etc., a universe like ours. On the other hand, in the no–boundary proposal, there is simply no defined initial state. Hawking says, "In the usual approach it is hot Big–Bang cosmology. The hypothetical particle or field thought to be responsible for inflation is called the inflaton field.

Inflation suggests that there was a period of exponential expansion in the very early universe. The expansion is exponential because the distance between any two fixed observers is increasing exponentially, due to the metric expansion of space (a space–time with this property is called a de Sitter space, see below). The physical conditions from one moment to the next are stable: the rate of expansion, called the Hubble parameter, is nearly constant, which leads to high levels of symmetry. Inflation is often called a period of accelerated expansion because the distance between two fixed observers is increasing at an accelerating rate as they move apart (however, this does not mean that the Hubble parameter is increasing).

Cosmic inflation has the important effect of smoothing out inhomogeneities, anisotropies and the curvature of space. This pushes the universe into a very simple state, in which it is completely dominated by the inflaton field and the only significant inhomogeneities are the tiny quantum fluctuations in the inflaton. Inflation also dilutes exotic heavy particles, such as the magnetic monopoles predicted by many extensions to the Standard Model of particle physics. If the universe was only hot enough to form such particles before a period of inflation, they would not be observed in nature, as they would be so rare that it is quite likely that there are none in the observable universe. Together, these effects are called the inflationary ‘no–hair Theorem’ by analogy with Hawking’s no hair Theorem for black holes. The ‘no–hair’ Theorem works essentially because the universe expands by an enormous factor during inflation. In an expanding universe, energy densities generally fall as the volume of the universe increases. For example, the density of ordinary ‘cold’ matter (dust) goes as the inverse of the volume: when linear dimensions double, the energy density goes down by a factor of eight. The energy density in radiation goes down even more rapidly as the universe expands: when linear dimensions are doubled, the energy density in radiation falls by a factor of sixteen. During inflation, the energy density in the inflaton field is roughly constant. However, the energy density in inhomogeneities, curvature, anisotropies and exotic particles is falling, and through sufficient inflation these become negligible. This leaves an empty, flat, and symmetric universe, which is filled with radiation when inflation ends.

A key requirement is that inflation must continue long enough to produce the present observable universe from a single, small inflationary Hubble volume. This is necessary to ensure that the universe appears flat, homogeneous and isotropic at the largest observable scales. This requirement is generally thought to be satisfied if the universe expanded by a factor of at least \(10^{56}\) during inflation. At the end of inflation, a process called reheating occurs, in which the inflaton particles decay into the radiation that starts the hot Big–Bang. It is not known how long inflation lasted but it is usually thought to be extremely short compared to the age of the universe. Assuming that the energy scale of inflation is between \(10^{15}\) and \(10^{16}\) GeV, as is suggested by the simplest models (like the Friedmann equation, see section on Cosmological Dynamics below), the period of inflation responsible for the observable universe probably lasted roughly \(10^{-33}\) seconds.
difficult to explain how inflation began. But it occurs naturally in top-down approach with no-boundary condition. It doesn’t need fine tuning.”

To do top-down cosmology, Hawking and Hertog first take a whole raft of possible histories (i.e., evolving geometries), all of which would result in a universe with features familiar to us. “We then calculate the probability for other features of the universe, given the constraints. Top-down cosmology does not predict that all possible universes have to begin with a period of inflation, but that inflation occurs naturally within a certain subclass of universes,” Hertog says. The process creates a probability for each scenario, and so Hertog can see which kind of history is most likely. “What we find is that the inflating histories generally have the largest probability.”

String Theory Landscape

Hawking and Hertog’s top-down/no-boundary cosmology adds an interesting twist to the ongoing debate in physics about the existence of multiple universes. At issue is the fact that string theory, physicists most popular candidate for a theory of everything, describes not just one universe but a near infinity of them. Some physicists are willing to accept that these theoretical universes actually exist, both because string theory does not seem to favour any particular universe over all others, and because their existence could explain the apparently fine-tuned features of our universe. Hawking’s view is that the so-called string theory landscape is populated by the set of all possible histories. Rather than a branching set of individual universes, every possible version of a single universe exists simultaneously in a state of quantum superposition. When we choose to make a measurement, we select from this landscape a subset of histories that share the specific features measured. From observer’s perspective, the history of the universe is derived from that subset of histories. In other words, we choose our past.

4.3.2 Theories of Everything, Anthropic Principle and Wave Function of the Universe

If a cat, a cannonball, and an economics textbook are all dropped from the same height, they fall to the ground with exactly the same acceleration (9.8m/s²) under the influence of gravity. This equality of the gravitational accelerations of different things is one of the most accurately tested laws of physics. The accuracy record holder at the moment
is the lunar laser ranging demonstration that the Earth and the Moon fall with the same acceleration toward the Sun [Anderson and Williams (2001)]. The accelerations are known to be equal to an accuracy of a few parts in a thousand billion. The equality of gravitational accelerations of different things is an example of a regularity of nature. Everything falls in exactly the same way. The regularity is universal. No exceptions!

Identifying and explaining the regularities of nature is the goal of science. Physics, like other sciences, is concerned with the regularities exhibited by particular systems. Stars, atoms, fluid flows, high temperature superconductors, black holes, and the elementary particles are just some of the many examples. Studies of these specific systems define the various subfields of physics—astrophysics, atomic physics, fluid mechanics, and so forth. But beyond the regularities exhibited by specific systems, physics has a special charge. This is to find the laws that govern the regularities that are exhibited by all physical systems, without exception, without qualification, and without approximation. The equality of gravitational accelerations of different things is an example. These are usually called the fundamental laws of physics. Taken together they are called informally a theory of everything. S. Hawking has been a leader in the quest for these universal laws [Hawking (1984b)].

Ideas for the nature of the fundamental laws of nature have changed as experiment and observation have revealed new realms of phenomena and reached new levels of precision in the last century. But since they have been studied, it has been thought that the fundamental laws consist of two parts:

- The dynamical laws that govern regularities in time. Newton’s laws of motion governing the orderly progression of the planets, or the trajectory of a tennis ball are examples, as is the law that different things fall with the same acceleration in a gravitational field and the Einstein equation governing the evolution of the universe.
- The initial conditions that govern how things started out and therefore most often specify regularities in space.

This so-called Newtonian determinism, famously propagated by P.S. Laplace [Laplace (1951)], was the first theory of everything dated circa 1820.

Both parts of a theory of everything are needed to make any predictions. Newton’s dynamical laws by themselves do not predict the trajectory of
Quantum Leap

a tennis ball we might throw. To predict where it goes, we must also specify the position from which we throw it, the direction, and how fast. Technically, we must specify the ball’s *initial condition*. One of Hawking’s most famous achievements is such an initial condition [Hawking (1984b)], but not for tennis balls. Hawking’s *no–boundary initial condition* is for the whole universe [Hartle (2003b)].

The search for a theory of the dynamical laws has been seriously under way since the time of Newton. Classical mechanics, Newtonian gravity, Maxwell’s electrodynamics, special and general relativity, quantum mechanics, quantum field theory, and superstring theory are but some of the milestones in this search. But the search for a theory of the initial condition of the universe has been seriously under way for only the twenty years since Hawking’s pioneering work on the subject. Why this difference? The examples used above to discuss regularities governed by the two parts of a theory of everything hint at the answer. The trajectory of a tennis ball was used to illustrate the regularities of dynamical laws, and the large scale distribution of galaxies was used to illustrate the regularities implied by the law of the initial condition. There is a difference in the *kind* and *scale* of regularities that the two laws predict.

*Dynamical laws predict regularities in time.* It is a fortunate empirical fact that the fundamental dynamical laws are *local* — both in space and time. As Einstein said, physics is simple only locally. The trajectory of a tennis ball depends only on conditions that are nearby both in space and time, and not, for example, either on what is going on in distant parts of the universe or a long time ago. This is fortunate because that means that dynamical laws can be discovered and studied in laboratories on Earth and extrapolated, assuming locality, to the rest of the universe. For example, because it is local, the law that different things fall with the same acceleration in a gravitational field can be discovered by experiments in laboratories here, and indeed all over the galaxy. Without that simplicity of the dynamical laws in the here and now it is possible that we would never have discovered them [Hartle (2003b)].

By contrast, the regularities governed by the law of the initial condition of the universe occur on large, cosmological scales. The universe is not simple on small spatial scales. Look at the disorder or complexity in the room we’re in right now for example. But the universe is simple on large, cosmological scales — more or less the same in one direction as in any other, more or less the same in one place as in any other.
4.3.2.1 Quantum Mechanics

There is another way in which our vision of the fundamental laws and the nature of a theory of everything has changed since the times of Newton and Laplace. That is quantum mechanics. We do not yet know the final form the fundamental laws will take. But the inference is inescapable from the physics of the last seventy–five years that they will conform to that subtle framework of prediction we call quantum mechanics.

Recall that in quantum mechanics, any system (the universe included) is described by a wave function $\Psi = \Psi(t)$. There is a dynamical law called the Schrödinger equation which governs how the wave function changes in time,

$$i\hbar \frac{d\Psi(t)}{dt} = H\Psi(t) \quad \text{(quantum dynamical law)}.$$

Here the operator $H$, called the Hamiltonian, summarizes the dynamical theory. There are different forms of $H$ for Maxwell’s electrodynamics, for a theory of the strong interactions, etc. Like Newton’s laws of motion, the Schrödinger equation does not make any predictions by itself, it requires an initial condition. This is $\Psi(0)$, (initial condition).

When we consider the universe as a quantum–mechanical system, this initial condition is Hawking’s no–boundary quantum wave function of the universe \cite{Hawking (1984b)}

Probabilities are the key difference between classical and quantum mechanics. Let’s first think about probabilities in classical physics. If I say that there is a 60% chance of hitting an audience member if I toss a ball in this room, I am not expressing a lack of confidence that its trajectory will be governed by the deterministic laws of Newtonian mechanics. Rather the 60% reflects my ignorance of the exact initial speed I’ll impart to the ball, of the influence of air on its motion, and perhaps my ability to do an accurate calculation. If I practice to control the initial condition when it’s thrown, the subsequent evolution of the tennis ball becomes more certain. Probabilities in classical physics result from ignorance.

But in quantum mechanics probabilities are fundamental and uncertainty inevitable. No amount of careful determination of the present state of the tennis ball will achieve certainty for its trajectory. In quantum mechanics there is some probability that a ball will take any trajectory as it leaves my hand. However, in classical situations one trajectory (the one...
obeying Newton’s laws) is much more probable than all the others. The determinism of classical physics is an approximation, but an approximation on which we can rely in many practical circumstances [Hartle (2003b)].

4.3.2.2 Scientific Reduction

Where do all the other regularities in the universe come from, those particular to specific systems — those of the behavior of cats as they fall, those studied by the environmental sciences like biology, geology, economics, and psychology? They are the results of chance events that occur naturally over the history of a quantum mechanical universe. As M. Gell-Mann puts it in [Gell-Mann (1994)], they are frozen accidents. “Chance events of which particular outcomes have a multiplicity of long–term consequences, all related by their common ancestry”.

The regularities of cats probably do depend a little bit on the fundamental physical laws, for example, an initial condition that is smooth across the universe, leads to three spatial dimensions, etc. But the origin of most of their regularities can be traced to the chance events of four billion years of biological evolution. Cats behave in similar ways because they have a common ancestry and develop in similar environments. The mechanisms which produce those chance events that led to cats are very much dependent on fundamental biochemistry and ultimately atomic physics. But the particular outcomes of those chances have little to do with the theory of everything [Hartle (2003b)].

Do psychology, economics, biology reduce to physics? The answer is YES, because everything considered in those subjects must obey the universal, fundamental laws of physics. Every one of the subjects of study in these sciences, humans, market tables, historical documents, bacteria, cats, etc., fall with the same acceleration in a gravitational field. The answer is NO, because the regularities of interest in these subjects are not predicted by the universal laws with near certainty even in principle. They are frozen quantum accidents that produce emergent regularities. The answer depends upon what we mean by reduce.

In summary:

- The fundamental laws of physics constituting a ‘theory of everything’ are those which specify the regularities exhibited by every physical system, without exception, without qualification, and without approximation.
- A theory of everything is not (and cannot be) a theory of everything
in a quantum–mechanical universe.

- The regularities of human history, personal psychology, economics, biology, geology, etc. are consistent with the fundamental laws of physics, but do not follow from them.

But remember also, especially on this occasion, that all the beautiful regularities that we observe in the universe, certain or not, predictable or not, could be the result of quantum chances following from the fundamental dynamical theory and Hawking’s no-boundary wave function of the universe. For more details on theories of everything, see [Hartle (2003b)].

4.3.2.3 Anthropic Reasoning in Quantum Cosmology

Prediction in quantum cosmology requires a specification of the universe’s quantum dynamics and its (initial) quantum state. We expect only a few general features of the universe to be predicted with probabilities near unity conditioned on the dynamics and quantum state alone. Most useful predictions are of conditional probabilities that assume additional information beyond the dynamics and quantum state. Anthropic reasoning utilizes probabilities conditioned on ‘us’. In this section, following [Hartle (2004b)], we discuss anthropic reasoning and quantum cosmology.

If the universe is a quantum mechanical system, then it has a quantum state. This state provides the initial condition for cosmology. A theory of this state is an essential part of any final theory summarizing the regularities exhibited universally by all physical systems and is the objective of the subject of quantum cosmology. This section is concerned with the role the state of the universe plays in anthropic reasoning, that is the process of explaining features of our universe from our existence in it [Barrow and Tipler (1986)]. The main idea is that anthropic reasoning in a quantum mechanical context depends crucially on assumptions about the universe’s quantum state.

A Model Quantum Universe

Every prediction in a quantum mechanical universe depends on its state if only very weakly. Quantum mechanics predicts probabilities for alternative possibilities, most generally the probabilities for alternative histories of the universe. The computation of these probabilities requires both a theory of the quantum state as well as the theory of the dynamics specifying its evolution.
To make this idea concrete while keeping the discussion manageable, we consider a model quantum universe. The details of this model are not essential to the subsequent discussion of anthropic reasoning but help to fix the notation for probabilities and provide a specific example of what they mean. Particles and fields move in a large, perhaps expanding box, say presently 20,000 Mpc on a side. Quantum gravity is neglected, which is an excellent approximation for accessible alternatives in our universe later than $10^{-43}$ s from the Big–Bang. Spacetime geometry is thus fixed with a well defined notion of time and the usual quantum apparatus of Hilbert space, states, and their unitary evolution governed by a Hamiltonian can be applied [Hartle (2004b)].

The Hamiltonian $H$ and the state $|\Psi\rangle$ in the Heisenberg picture are the assumed theoretical inputs to the prediction of quantum mechanical probabilities. Alternative possibilities at one moment of time $t$ can be reduced to yes/no alternatives represented by an exhaustive set of orthogonal projection operators $\{P_\alpha(t)\}, \alpha = 1, 2, \ldots$ in this Heisenberg picture. The operators representing the same alternatives at different times are connected by

$$P_\alpha(t) = e^{iHt/\hbar}P_\alpha(0)e^{-iHt/\hbar}. \quad (4.40)$$

For instance, the $P$’s could be projections onto an exhaustive set of exclusive ranges of the center-of-mass position of the Earth labelled by $\alpha$. The probabilities $p(\alpha)$ that the Earth is located in one or another of these regions at time $t$ is

$$p(\alpha|H, \Psi) = \|P_\alpha(t)|\Psi\rangle\|^2. \quad (4.41)$$

The probabilities for the Earth’s location at a different time is given by the same formula with different $P$’s computed from the Hamiltonian by (4.40). The notation $p(\alpha|H, \Psi)$ departs from usual conventions (e.g., Hartle (1993)) to indicate explicitly that all probabilities are conditioned on the theory of the Hamiltonian $H$ and quantum state $|\Psi\rangle$.

Most generally quantum theory predicts the probabilities of sequences of alternatives at a series of times, that is quantum histories. An example is a sequence of ranges of center of mass position of the Earth at a series of times giving a coarse–grained description of its orbit. Sequences of sets of alternatives $\{P_{\alpha_k}(t_k)\}$ at a series of times $t_k, k = 1, \ldots, n$ specify a set of alternative histories of the model universe. An individual history $\alpha$ in the set corresponds to a particular sequence of alternatives $\alpha \equiv (\alpha_1, \alpha_2, \ldots, \alpha_n)$
and is represented by the corresponding chain of projection operators \( C_\alpha \) \cite{Hartle2004b}

\[
C_\alpha \equiv P_{n_\alpha}^{(t_n)}(t_n) \cdots P_{1_\alpha}^{(t_1)}(t_1) \ , \ \alpha \equiv (\alpha_1, \ldots, \alpha_n).
\]

(4.42)

The probabilities of the histories in the set are given by

\[
p(\alpha|H, \Psi) \equiv p(\alpha_n, \ldots, \alpha_1|H, \Psi) = \|C_\alpha|\Psi\rangle\|^2,
\]

(4.43)

provided the set decoheres, i.e., provided the branch state vectors \( C_\alpha|\Psi\rangle \) are mutually orthogonal. Decoherence ensures the consistency of the probabilities (4.43) with the usual rules of probability theory.\footnote{For a short introduction to decoherence see \cite{Hartle1993} or any of the classic expositions of decoherent (consistent) histories quantum theory \cite{Griffiths2002, Omnès1996, Gell-Mann1994}.}

To use either (4.41) or (4.43) to make predictions, a theory of both \( H \) and \( |\Psi\rangle \) is needed. No state; no predictions.

**What is Predicted?**

Once M. Gell-Mann asked J. Hartle, “If you know the wave function of the universe, why are not you rich?” (see \cite{Hartle2004b}). Hartle’s answer was that there were unlikely to be any alternatives relevant to making money that were predicted as sure bets conditioned just on the Hamiltonian and quantum state alone. A probability \( p(\text{rise}|H, \Psi) \) for the stock market to rise tomorrow could be predicted from \( H \) and \( |\Psi\rangle \) through (4.41) in principle. But it seems likely that the result would be a useless \( p(\text{rise}|H, \Psi) \approx 1/2 \) conditioned just on the ‘no boundary’ wave function.

It’s plausible that this is the generic situation. To be manageable and discoverable, the theories of dynamics and the quantum state must be short, that is describable in terms of a few fundamental equations and the explanations of the symbols they contain. It’s therefore unlikely that \( H \) and \( |\Psi\rangle \) contain enough information to determine most of the interesting complexity of the present universe with significant probability \cite{Hartle1996, Hartle2003a}. We hope that the Hamiltonian and the quantum state are sufficient conditions to predict certain large scale features of the universe with significant probability. Approximately classical space–time, the number of large spatial dimensions, the approximate homogeneity and isotropy on scales above several hundred Mpc, and the spectrum of density fluctuations that were the input to inflaton are some examples of these. But even
a simple feature like the time the Sun will rise tomorrow will not be usefully predicted by our present theories of dynamics and the quantum state alone.

The time of sunrise does become predictable with high probability if a few previous positions and orientations of the Earth in its orbit are supplied in addition to $H$ and $|\Psi\rangle$. That is a particular case of a conditional probability of the form

$$p(\alpha|\beta, H, \Psi) = \frac{p(\alpha, \beta|H, \Psi)}{p(\beta|H, \Psi)},$$

(4.44)

for alternatives $\alpha$ (e.g., the times of sunrise) given $H$ and $|\Psi\rangle$ and further alternatives $\beta$ (e.g., a few earlier positions and orientations of the Earth). The joint probabilities on the right hand side of (4.61) are computed using (4.43) as described above.

Conditioning probabilities on specific information can weaken their dependence on $H$ and $|\Psi\rangle$ but does not eliminate it. That is because any specific information available to us as human observers (like a few positions of the Earth) is but a small part of that needed to specify the state of the universe. The $P_\beta$ used to define the joint probabilities in (4.61) by (4.43) therefore spans a very large subspace of Hilbert space. As a consequence $P_\beta|\Psi\rangle$ depends strongly on $|\Psi\rangle$. For example, to extrapolate present data on the Earth to its position 24 hours from now requires that the probability be high that it moves on a classical orbit in that time and that the probability be low that it is destroyed by a neutron star now racing across the galaxy at near light speed. Both of these probabilities depend crucially, if weakly, on the nature of the quantum state [Hartle (1994b)].

Many useful predictions in physics are of conditional probabilities of the kind discussed in this section. We next turn to the question of whether we should be part of the conditions.

**Anthropic Reasoning**

In calculating the conditional probabilities for predicting some of our observations given others, there can be no objection of principle to including a description of ‘us’ as part of the conditions [Hartle (2004b)],

$$p(\alpha|\beta, ‘us’, H, \Psi).$$

(4.45)

Drawing inferences using such probabilities is called *anthropic reasoning*. The motivation is the idea that probabilities for certain features of the universe might be sensitive to this inclusion.
The utility of anthropic reasoning depends on how sensitive probabilities like (4.63) are to the inclusion of ‘us’. To make this concrete, consider the probabilities for a hypothetical cosmological parameter we will call Λ. We will assume that $H$ and $|\Psi\rangle$ imply that Λ is constant over the visible universe, but only supply probabilities for the various constant values it might take through (4.43). We seek to compare $p(\Lambda|H,\Psi)$ with $p(\Lambda|\text{us}',H,\Psi)$. In principle, both are calculable from (4.43) and (4.61). There are three possible ways they might be related [Hartle (2004b)]:

- $p(\Lambda|H,\Psi)$ is peaked around one value. The parameter Λ is determined either by $H$ or $|\Psi\rangle$, or by both. Anthropic reasoning is not necessary; the parameter is already determined by fundamental physics.
- $p(\Lambda|H,\Psi)$ is distributed and $p(\Lambda|\text{us}',H,\Psi)$ is also distributed. Anthropic reasoning is inconclusive. One might as well measure the value of Λ and use this as a condition for making further predictions, i.e., work with probabilities of the form $p(\alpha|\Lambda,H,\Psi)$.
- $p(\Lambda|H,\Psi)$ is distributed but $p(\Lambda|\text{us}',H,\Psi)$ is peaked. Anthropic reasoning helps to explain the value of Λ.

The important point to emphasize is that a theoretical hypothesis for $H$ and $|\Psi\rangle$ is needed to carry out anthropic reasoning. Put differently, a theoretical context is needed to decide whether a parameter like Λ can vary, and to find out how it varies, before using anthropic reasoning to restrict its range. The Hamiltonian and quantum state provide this context. Below we will consider the situation where the state is imperfectly known.

While there can be no objections of principle to including ‘us’ as a condition for the probabilities of our observations, there are formidable obstacles of practice [Hartle (2004b)]:

- We are complex physical systems requiring an extensive environment and a long evolutionary history whose description in terms of the fundamental variables of $H$ and $|\Psi\rangle$ may be uncertain, long, and complicated.
- The complexity of the description of a condition including ‘us’ may make the calculation of the probabilities long or impossible as a practical matter.

In practice, therefore, anthropic probabilities (4.63) can only be estimated or guessed. Theoretical uncertainty in the results is thereby introduced.
The objectivity striven for in physics consists, at least in part, in using probabilities that are not too sensitive to ‘us’. We would not have science if anthropic probabilities for observation depended significantly on which individual human being was part of the conditions. The existence of schizophrenic delusions shows that this is possible so that the notion of ‘us’ should be restricted to exclude such cases.

For these reasons it is prudent to condition probabilities, not on a detailed description of ‘us’, but on the weakest condition consistent with ‘us’ that plausibly provides useful results. A short list of conditions of roughly decreasing complexity might include [Hartle (2004b)]:

- human beings;
- carbon-based life;
- information gathering and utilizing systems (IGUSes);
- at least one galaxy;
- a universe older than 5 Gyr;
- no condition at all.

For example, the probabilities used to bound the cosmological constant \( \Lambda \) [Barrow and Tipler (1986); Weinberg (2005)] make use of the fourth and fifth on this list under the assumption that including earlier ones will not much affect the anthropically-allowed range for \( \Lambda \). To move down in the above list of conditions is to move in the direction of increasing theoretical certainty and decreasing computational complexity. With anthropic reasoning, less is more.

**Ignorance is NOT Bliss**

The quantum state of a single isolated subsystem generally cannot be determined from a measurement carried out on it. That is because the outcomes of measurements are distributed probabilistically and the outcome of a single trial does not determine the distribution. Neither can the state be determined from a series of measurements because measurements disturb the state of the subsystem. The Hamiltonian can not be inferred from a sequence of measurements on one subsystem for similar reasons. In the same way, we cannot generally determine either the Hamiltonian or the quantum state of the universe from our observations of it. Rather these two parts of a final theory are theoretical proposals, inferred from partial data to be sure, but incorporating theoretical assumptions of simplicity, beauty, coherence, mathematical precision, etc. To test these proposals we
search among the conditional probabilities they imply for predictions of observations yet to be made with probabilities very near one. When such predictions occur we count it a success of the theory, when they do not we reject it and propose another. The main question is, do we need a theory of the quantum state? To analyze this question, let us consider various degrees of theoretical uncertainty about it [Hartle (2004b)].

**Total Ignorance.** In the model cosmology in a box of Section II, theoretical uncertainty about the quantum state can be represented by a density matrix \( \rho \) that specifies probabilities for its eigenstates to be \( |\Psi\rangle \).

Total ignorance of the quantum state is represented by a \( \rho \) proportional to the unit matrix. To illustrate this and the subsequent discussion, assume for the moment that the dimension of the Hilbert space is very large but finite. Then total ignorance of the quantum state is represented by [Hartle (2004b)]

\[
\rho_{\text{tot. ign.}} = \frac{I}{Tr(I)},
\]

which assigns equal probability to any member of any complete set of orthogonal states.

The density matrix (4.69) predicts thermal equilibrium, infinite temperature, infinitely large field fluctuations, and maximum entropy [Hartle (2003a)]. In short, its predictions are inconsistent with observations. This is a more precise way of saying that every useful prediction depends in some way on a theory of the quantum state. Ignorance is never bliss.

**What We Know.** A more refined approach to avoiding theories of the quantum state is to assume that it is unknown except for reproducing our present observations of the universe. The relevant density matrix is [Hartle (2004b)]

\[
\rho_{\text{obs}} = \frac{P_{\text{obs}}}{Tr(P_{\text{obs}})},
\]

where \( P_{\text{obs}} \) is the projection on our current observations, that is “what we know”. “Observations” in this context mean what we directly observe and record here on Earth and not the inferences we draw from this data about the larger universe. That is because those inferences are based on assumptions about the very quantum state that (4.47) aims to ignore. For instance, we observed nebulae long before we understood what they were or where they are. The inference that the nebulae are distant clusters of stars and gas relies on assumptions about how the universe is structured.
on very large scales that are in effect weak assumptions on the quantum state.

Even if we made the overly generous assumption that we had somehow directly observed and recorded every detail of the volume 1 km above the surface of the Earth, say at a 1 mm resolution, that is still a tiny fraction ($\sim 10^{-60}$) of the volume inside the present cosmological horizon. The projection operator $P_{\text{obs}}$ therefore defines a very large subspace of Hilbert space. We can expect that the entropy of the density matrix (4.47) will therefore be near maximal, close to that of (4.69), and its predictions similarly inconsistent with further observations.

In the context of anthropic reasoning, these results show that conditioning probabilities on ‘us’ alone is not enough to make useful predictions. Rather, a theory of $H$ and $|\Psi\rangle$ are needed in addition as described in the previous section.

Let us hope that one day we will have a unified theory based on a principle that will specify both quantum dynamics ($H$) and a unique quantum state of the universe ($|\Psi\rangle$). That would truly be a final theory and a proper context for anthropic reasoning. For more details, see [Hartle (2004b)].

4.3.2.4 Quantum State of the Universe

The central question of quantum cosmology is, What is the quantum state of the universe? To ask this question is to assume that the universe is a quantum mechanical system. We perhaps have little direct evidence of peculiarly quantum mechanical phenomena on large and even familiar scales, but there is no evidence that the phenomena that we do see cannot be described in quantum mechanical terms and explained by quantum mechanical laws. Further, every major candidate for a fundamental dynamical law from the standard model to M–theory conforms to the quantum mechanical framework for prediction. If this framework applies to the whole thing, there must be a quantum state of the universe [Hartle (2003a)].

Final Theories

The final theory predicts the regularities that are exhibited by all physical systems, without exception, without qualification, and without approximation. A possible view at present is that it consists of two parts [Hartle (2003a)]:

• A universal dynamical law such as string theory or its successors;
A law for the initial quantum state of the universe such as Hawking’s no-boundary wave function of the universe.

In a model universe in a box these two parts are represented by the Hamiltonian specifying the form of the Schrödinger equation

$$i\hbar \frac{d|\Psi(t)\rangle}{dt} = H |\Psi(t)\rangle$$

(4.48)

and the initial quantum state

$$|\Psi(0)\rangle.$$  \hspace{1cm} (4.49)

Both of these pieces are necessary for prediction. The Schrödinger equation makes no predictions by itself. The probabilities $p_\alpha$ predicted by quantum mechanics for a set of alternatives represented by projection operators $\{P_\alpha\}$ are

$$p_\alpha = ||P_\alpha |\Psi(t)\rangle||^2.$$  \hspace{1cm} (4.50)

To compute these the quantum state is needed at least at one time. No state, no predictions.

To put this in a different way, if the state is arbitrary, the predictions are arbitrary. Pick any probabilities $p_\alpha$ you like for the alternatives $P_\alpha$. There is some state that will reproduce them. For example, we can take

$$|\Psi(t)\rangle = \sum_\alpha p_\alpha^{\frac{1}{2}} |\Psi_\alpha\rangle$$

(4.51)

where the $|\Psi_\alpha\rangle$ are any set of eigenstates of the $P_\alpha$’s

$$P_\alpha |\Psi_\beta\rangle = \delta_{\alpha\beta} |\Psi_\beta\rangle.$$  \hspace{1cm} (4.52)

The $|\Psi(t)\rangle$ constructed according to (4.51) will reproduce the pre-assigned probabilities $p_\alpha$ \cite{Hartle2003a}.

Neither is ignorance bliss. If we assume we know nothing about the state of the universe in a box then we should make predictions with a density matrix proportional to unity

$$\rho = \frac{1}{Tr(I)},$$

(4.53)

reflecting that ignorance. But this density matrix corresponds to equilibrium at infinite temperature and its predictions are nothing like the universe we live in. In particular, there would be no evolution since $[H, \rho] = 0$. There would be no second law of thermodynamics since the entropy $-Tr[\rho \log \rho]$ is
already at its maximum possible value. There would be no classical behavior since, although the expected value of a field averaged over a space–time volume $R$ might be finite, its fluctuations, $\langle \phi(R)^2 \rangle$, would be infinite [Hartle (2003a)].

The search for a unified fundamental dynamical law has been seriously under way at least since the time of Newton, with string theory or its generalizations being the most actively investigated direction today. By contrast, the search for a theory of the quantum state of the universe has only been actively under way since the time of Hawking. Why this difference?

Dynamical laws govern regularities in time and it is an empirical fact that the basic dynamical laws are local in space on scales above the Planck length. The laws that govern regularities in time across the whole universe are therefore discoverable and testable in laboratories on earth. By contrast, many of the regularities predicted with near certainty by the quantum state of the universe are mostly in space on large cosmological scales. Only recently has there been enough data to confront theory with observation. That difference in the nature of the predicted regularities, or their difference in scales, should not obscure the fact that the state is just as much a part of the final theory as is its Hamiltonian.

Given these differences, what grounds do we have to hope that we can discover the quantum state of the universe? There are two: The first is the simplicity of the early universe revealed by observation — more homogeneous, more isotropic, more nearly in thermal equilibrium than the universe is today. It is therefore possible that the universe has a simple, discoverable initial quantum state and that all of the complex universe of galaxies, stars, planets, and life today arose from quantum accidents that have happened since and the action of gravitational attraction. The second reason is the idea that the quantum state and the dynamical theory may be naturally connected as in Hawking’s no–boundary theory [Hartle (2003a)].

Effective Theories

We are used to the idea of effective dynamical theories that accurately describe limited ranges of phenomena. The Navier–Stokes equations, non–relativistic quantum mechanics, general relativity, quantum electrodynamics, and the standard model of particle physics are all familiar examples. To construct an effective theory we typically assume a coarse–grained description (restricting attention to energies below the Planck scale for instance) and assume some simple property that the state might predict there (clas-
Cosmology too has its effective theory and its standard model. This is summarized neatly by the following list (extended) from Rees (2001):

- space–time is classical, governed by the Einstein equation
- our universe is expanding,
- from a hot Big–Bang,
- in which light elements were synthesized,
- there was a period of inflation,
- which led to a flat universe today,
- structure was seeded by gaussian irregularities,
- which are relics of quantum fluctuations,
- the dominant matter is cold and dark,
- but a cosmological constant (or quintessence) is dynamically dominant.

Possibly all current observations in cosmology, at least the large scale ones, can be compressed into an effective ‘Standard model’ based on this list of ten assumptions and a few cosmological parameters. That is not unlike the situation in particle physics where most observations can be compressed into the Lagrangian of the Standard model and its eighteen or so parameters.

However, the success of such effective theories which operate in limited ranges of phenomena should not obscure the need to find fundamental ones which apply to all phenomena without qualification and without approximation. It would be inconsistent, we believe, to pursue a fundamental dynamical theory in the face of a successful effective, standard dynamical model, and not pursue a fundamental theory of the state of the universe because of the success of its effective, standard cosmological model. That not least because the fundamental theory could provide a unified explanation of its assumptions.

It must be said, however, that when the natural domains of fundamental theories are as far from controllable experiments as string theory and the quantum initial condition the possibility of definitive tests seems to recede. It could be that the predictions of string theory are limited to general relativity, gauge theories, supersymmetry, and the parameters of the standard particle model. In a similar way the predictions of the state of the universe could be limited to classical space–time, the initial conditions for inflation, and the quantum fluctuations that satisfy large scale structure. Perhaps that is prediction enough Hartle (2003a).
Directions

Continuing the search for a final theory incorporating dynamics and the initial quantum state is certainly one direction. But we would like to mention three questions that might lead to different approaches to the main one [Hartle (2003a)].

What’s Environmental? Which features of the observed universe follow entirely from the dynamical theory (H) and which follow entirely from the initial condition (|Ψ(0)⟩), and which are the result of quantum accidents that occurred over the course of the universe’s history with probabilities specified by the combination of H and |Ψ(0)⟩. Those that depend significantly on |Ψ(0)⟩ are called ‘environmental’. Some version of this question was number one on the list of top ten questions for the next millennium prepared by string theorists at the Strings 2000 conference [Duff et al. (2001)].

Take the coupling constants in effective dynamical theories for instance. The viscosities and equation–of–state in the Navier–Stokes equation are certainly environmental. They vary with system, place, and time. But at a given energy do the coupling constants of the standard model of the elementary particle interactions vary with place and time or with the possible history of the universe? If so then the initial quantum state is central to determining their probabilities.

Why Quantum Mechanics? The founders of quantum mechanics thought that the inherent indeterminacy of quantum theory “reflected the unavoidable interference in measurement dictated by the magnitude of the quantum of the action” (Bohr). But why then do we live in a quantum mechanical universe which, by definition, is never measured from the outside?

The most striking general feature of quantum mechanics is its exact linearity, the principle of superposition. But why should there be a principle of superposition in quantum cosmology which has only a single quantum state?

Why a Division into Dynamics and Initial Condition? The schema for a final theory which I have been describing posits a separate theory of dynamics and quantum state. Could they be connected? They already are in Hawking’s no–boundary wave function [Hawking (1984b)]

\[ \Psi = \int D[g] D[\phi] e^{-I[g,\phi]}, \] (4.54)

where the action for metric \(g_{\alpha\beta}(x)\) coupled to matter \(\phi(x)\) determines both
the state and quantum dynamics. Is there a principle that determines both? Is there a connection between superstring theory and its successors and a unique quantum state?

A unified quantum theory of state and dynamics would be truly a final theory. Pursuing that vision is surely a direction for theoretical physics and cosmology [Hartle (2003a)].

4.3.2.5 From Quantum Mechanics to Quantum Gravity

How do our ideas about quantum mechanics affect our understanding of space–time? This familiar question leads to quantum gravity [Hartle (2006)]. This subsection addresses a complementary question: How do our ideas about space–time affect our understanding of quantum mechanics?

Familiar non–relativistic quantum theory illustrates how quantum mechanics incorporates assumptions about space–time. The Schrödinger equation governs the evolution of the state in between measurements

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi . \quad (4.55)$$

The state vector is ‘reduced’ at the time of a measurement according to the second law of evolution:

$$\psi \rightarrow P\psi/||P\psi||, \quad (4.56)$$

where $P$ is the projection on the outcome of the measurement. Both of these laws of evolution assume a fixed background space–time. A fixed geometry of space–time is needed to define both the $t$ in the Schrödinger equation and the space–like surface on which the state vector is reduced.

But, in quantum gravity, the geometry of space–time is not fixed. Rather geometry is a quantum variable, fluctuating and generally without a definite value. There is no fixed $t$. Quantum mechanics must therefore be generalized to deal with quantum space–time. This is sometimes called the ‘problem of time’ in quantum gravity.

Already our ideas about quantum theory have evolved as out ideas about space–time have changed. Milestones in the evolution of our concepts of space and time include: the separate space and absolute time of Newtonian physics, Minkowski space–time with different times in different Lorentz frames, the curved but fixed space–time of general relativity, the quantum fluctuations of space–time in quantum gravity, and the ideas of string/M-theory and loop quantum gravity that space–time is an approximation to something more fundamental. Changes in quantum theory have reflected
this evolution. Non-relativistic quantum mechanics incorporates Newtonian time in the Schrödinger equation and the second law of evolution. Any one of the possible time–like directions in Minkowski space can be used to describe the unitary evolution of quantum fields and the results of different choices are unitarily equivalent. Quantum field theories in curved space–times based on different foliations by space–like surfaces are not generally unitarily equivalent. In quantum gravity there is no fixed space–time through which a state can unitarily evolve. Quantum mechanics therefore needs to be generalized for quantum gravity so that it does not require a fixed space–time foliable by space–like surfaces. And, if space–time is not fundamental, quantum mechanics will certainly need to be generalized for whatever replaces it.

However, familiar quantum mechanics also needs to be generalized for cosmology. This generalization is needed so that quantum theory can apply to closed systems such as the universe as a whole containing both observers and observed, measuring apparatus and measured subsystems (if any). These two generalizations can be connected in a common framework called \textit{generalized quantum theory} which is abstracted from the consistent (or decoherent) histories formulation of the quantum mechanics of closed systems \cite{Griffiths2004}.

The principles of generalized quantum mechanics were introduced in \cite{Hartle1991} and developed more fully for example in \cite{Hartle1995}. The principles have been axiomatized in a rigorous mathematical setting by \cite{Isham1994}. Three elements are needed to specify a generalized quantum theory \cite{Hartle2006}:

(1) The sets of \textit{fine-grained histories}. These are the most refined possible description of a closed system.

(2) The allowed \textit{coarse grainings}. A coarse graining of a set of histories is generally a partition of that set into mutually exclusive classes \(\{c_\alpha\}, (\alpha \text{ discrete})\) called \textit{coarse-grained histories}. The set of classes constitutes a set of coarse-grained histories with each history labelled by the discrete index \(\alpha\).

(3) A \textit{decoherence functional} defined for each allowed set of coarse-grained histories which measures the interference between pairs of histories in the set and incorporates a theory of the initial condition and dynamics of the closed system. A decoherence functional \(D(\alpha', \alpha)\) must satisfy the following properties:
Quantum Universe

(i) **Hermiticity:**
\[ D(\alpha', \alpha) = D^\ast(\alpha, \alpha') \]

(ii) **Positivity:**
\[ D(\alpha, \alpha) \geq 0. \]

(iii) **Normalization:**
\[ \sum_{\alpha' \in \bar{\alpha}} D(\alpha', \alpha) = 1. \]

(iv) **The Principle of Superposition:** If \( \{\bar{c}_\alpha\} \) is a coarse graining of a set of histories \( \{c_\alpha\} \), that is, a further partition into classes \( \{\bar{c}_\alpha\} \), then
\[ D(\bar{\alpha}', \bar{\alpha}) = \sum_{\alpha' \in \bar{\alpha}'} \sum_{\alpha \in \bar{\alpha}} D(\alpha', \alpha). \]

Once these three elements are specified the process of prediction proceeds as follows: A set of histories is said to (medium) decohere if all the “off-diagonal” elements of \( D(\alpha', \alpha) \) are sufficiently small. The diagonal elements are the probabilities \( p(\alpha) \) of the individual histories in a decoherent set. These two definitions are summarized in the one relation
\[ D(\alpha', \alpha) \approx \delta_{\alpha'\alpha} p(\alpha). \quad (4.57) \]

As a consequence of (4.57) and properties (i)-(iv) above, the numbers \( p(\alpha) \) lie between zero and one, sum to one, and satisfy the most general form of the probability sum rules
\[ p(\bar{\alpha}) = \sum_{\alpha \in \bar{\alpha}} p(\alpha) \]
for any coarse graining \( \{\bar{c}_\alpha\} \) of the set \( \{c_\alpha\} \). The \( p(\alpha) \) are therefore probabilities. They are the predictions of generalized quantum mechanics for the possible coarse-grained histories of the closed system that arise from the theory of its initial condition and dynamics incorporated in the construction of \( D \).

Feynman’s 1948 space–time formulation of quantum mechanics [Feynman (1948)] supplies one route to constructing a fully 4D generalized quantum theory of space–time geometry. The quantum mechanics of a non–relativistic particle moving in one dimension \( (x = x(t)) \) between time \( t = 0 \) and time \( t = T \) provides the simplest example. The particle’s dynamics is assumed specified by an action functional \( S[x(t)] \) and its initial quantum state is assumed to be a particular state vector \( |\psi\rangle \).
(1) **Fine–grained histories**: These are all paths $x(t)$ between $t = 0$ and $t = T$.

(2) **Coarse–grainings**: An allowed coarse graining is any partition of the paths into an exhaustive set of exclusive classes $c_\alpha$, ($\alpha$ discrete), each class being a *coarse-grained history*. For instance, the paths could be partitions by specifying a set of spatial intervals $\Delta_i$, $i = 1, 2, \cdots$ and giving which two intervals $\alpha = (i,j)$ the particle passes through at two times. An example of a *space–time coarse graining* is provided by specifying a space–time region $R$ and partitioning the paths into the class $c_0$ which never pass through $R$ and the class $c_1$ that pass through $R$ sometime.

(3) **Decoherence functional**: In a given set of coarse-grained histories $\{c_\alpha\}$ construct *branch state vector* $|\psi_\alpha\rangle$ for each coarse grained history by summing $\exp(iS)$ over all the paths in $c_\alpha$ and applying that to the initial state $|\psi\rangle$, viz.

$$|\psi_\alpha\rangle \equiv \int_{c_\alpha} D[x] \exp\{iS[x(t)]/\hbar\}|\psi\rangle.$$ 

The decoherence functional is

$$D(\alpha', \alpha) = \langle \psi_{\alpha'}|\psi_\alpha\rangle.$$ 

This space–time formulation of non–relativistic quantum mechanics is easy to visualize, fully 4D, manifests Lagrangian symmetries, and has a close connection to the semiclassical approximation. It incorporates both unitary evolution and the reduction of the state vector in a unified way [Caves (1986)].

The space–time formulation is equivalent to usual Hamiltonian quantum mechanics when the fine grained histories are *single valued in a time* as in non–relativistic quantum mechanics and Minkowski space quantum field theory. This fully 4D formulation generalizes usual quantum mechanics when the histories do not have this property, for instance if there is no fixed time or the histories are not single valued in time. But in those cases we cannot expect to find a notion of state of the system at a moment of time or its unitary evolution through time.

Here are several model situations [Hartle (2006)]:

- Spacetime alternatives extended over time such as those defined by field averages over space–time regions with extent both in time and space [Hartle (1991b)].
• Time-neutral quantum mechanics without a quantum mechanical arrow of time but with both initial and final conditions [Hartle and Gell-Mann (1994)].

• Quantum field theory in fixed background space–times that are not foliable by space–like surfaces such as space–times with closed time–like curves, spacetimes exhibiting topology change, and evaporating black hole space–times [Hartle (1994c) Hartle (1998)].

• Histories that move backward in time such as those of a single relativistic particle moving in 4D flat space–time [Hartle (1995a)].

For each of these examples the three ingredients of generalized quantum theory were exhibited, fine grained histories, coarse graining, and decoherence functional.

Building on the lessons of these examples, a generalized quantum mechanics of quantum cosmological space–time geometry can be sketched. The fine grained histories are closed 4D cosmological metrics with 4D matter field configurations upon them. The allowed coarse grainings are partitions of these histories into 4D diffeomorphism invariant classes \( c_\alpha \). A decoherence functional \( D(\alpha', \alpha) \) is constructed using amplitudes defined by sums over the histories in the classes \( c'_\alpha \) and \( c_\alpha \), initial and final wave functions of the universe, and an inner product linking amplitudes and wave functions. The semiclassical limit for geometry is provided by the steepest descents approximations to the sums over metrics. What remains is a usual quantum field theory in the background space–time described by the metric which gives the biggest contribution to these sums. Thus, familiar familiar quantum mechanics is recovered for those initial conditions and those coarse-grainings in which space–time is fixed, classical, and can supply the necessary time for unitary evolution.

As a summary, we have [Hartle (2006)].

• Quantum mechanics can be generalized so that it is free from a fundamental notion of measurement, free of the need for a fixed background space–time, and free from the ‘problem of time’.

• General relativity as a theory of four-dimensional space–time is more general than its 3+1 initial value problem. Similarly, a fully four-dimensional formulation of quantum theory is more general than its 3+1 formulation in terms of states evolving unitarily through space–like surfaces in a fixed background space–time.

• In a 4D generalized quantum mechanics of space–time geometry there is
no ‘problem of time’, but there are also typically no states at a moment of time.

- In the context of a fully four-dimensional formulation of quantum theory, the familiar 3+1 quantum mechanics of states evolving unitarily through space–like surfaces is an approximation that is appropriate for those initial conditions and those coarse grained descriptions in which space–time geometry behaves classically and can supply the notion of time necessary to describe the evolution.

### 4.3.3 Quantum Gravity and Black Holes

In their search for quantum gravity, S. Hawking and R. Penrose use the straightforward application of quantum theory to general relativity \[\text{[Hawking and Israel (1979); Penrose (1967); Hawking and Penrose (1996)]}\], rather than following the more fashioned string theory approach (described below).

According to Hawking, “Einstein’s general relativity is a beautiful theory that agrees with every observation that has been made so far. It might require modifications on the Planck scale, and it might be only a low energy approximation to some more fundamental theory, like e.g., superstring theory, but it will not affect many of the predictions that can be get from gravity...” \[\text{[Hawking and Israel (1979)]}\].

**Space–Time Manifold, Gravity, Black Holes and Big–Bang**

The crucial technique for investigating Hawking–Penrose singularities and black holes, has been the study of the global causal structure of space–time \[\text{[Hawking and Israel (1979)]}\]. Define \(I^+(p)\) to be the set of all points of the space–time manifold \(M\) that can be reached from the point \(p\) by future directed time like curves. One can think of \(I^+(p)\) as the set of all events that can be influenced by what happens at \(p\). One now considers the boundary \(\hat{I}^+(S)\) of the future of a set \(S\). It is easy to see that this boundary cannot be time–like. For in that case, a point \(q\) just outside the boundary would be to the future of a point \(p\) just inside. Nor can the boundary of the future be space–like, except at the set \(S\) itself. For in that case every past directed curve from a point \(q\), just to the future of the boundary, would cross the boundary and leave the future of \(S\). That would be a contradiction with the fact that \(q\) is in the future of \(S\). Therefore, the boundary of the future is null apart from at \(S\) itself.
To show that each generator of the boundary of the future has a past end point on the set, one has to impose some global condition on the causal structure. The strongest and physically most important condition is that of global space–time hyperbolicity. The significance of global hyperbolicity for singularity theorems stems from the following (Hawking and Israel (1979) [Hawking and Penrose (1996)]. Let $U$ be globally hyperbolic and let $p$ and $q$ be points of $U$ that can be joined by a time like or null curve. Then there is a time–like or null–geodesic between $p$ and $q$ which maximizes the length of time like or null curves from $p$ to $q$. The method of proof is to show the space of all time like or null curves from $p$ to $q$ is compact in a certain topology. One then shows that the length of the curve is an upper semi–continuous function on this space. It must therefore attain its maximum and the curve of maximum length will be a geodesic because otherwise a small variation will give a longer curve.

One can now consider the second variation of the length of a geodesic $\gamma$. One can show that $\gamma$ can be varied to a longer curve if there is an infinitesimally neighboring geodesic from $p$ which intersects $\gamma$ again at a point $r$ between $p$ and $q$. The point $r$ is said to be conjugate to $p$. One can illustrate this by considering two points $p$ and $q$ on the surface of the Earth. Without loss of generality one can take $p$ to be at the north pole. Because the Earth has a positive definite metric rather than a Lorentzian one, there is a geodesic of minimal length, rather than a geodesic of maximum length. This minimal geodesic will be a line of longitude running from the north pole to the point $q$. But there will be another geodesic from $p$ to $q$ which runs down the back from the north pole to the south pole and then up to $q$. This geodesic contains a point conjugate to $p$ at the south pole where all the geodesics from $p$ intersect. Both geodesics from $p$ to $q$ are stationary points of the length under a small variation. But now in a positive definite metric the second variation of a geodesic containing a conjugate point can give a shorter curve from $p$ to $q$. Thus, on the Earth, the geodesic that goes down to the south pole and then comes up is not the shortest curve from $p$ to $q$.

The reason one gets conjugate points in space–time is that gravity is an

---

5Recall that an open set $U$ is said to be globally hyperbolic if:

1. For every pair of points $p$ and $q$ in $U$ the intersection of the future of $p$ and the past of $q$ has compact closure. In other words, it is a bounded diamond shaped region.

2. Strong causality holds on $U$. That is, there are no closed or almost closed time–like curves contained in $U$. 
attractive force. It therefore curves space–time in such a way that neighboring geodesics are bent towards each other rather than away. One can see this from the Newman–Penrose equation\(^6\)

\[
\frac{d\rho}{dv} = \rho^2 + \sigma^{ij} \sigma_{ij} + \frac{1}{n} R_{\alpha\beta} l^\alpha l^\beta, \quad (\alpha, \beta = 0, 1, 2, 3)
\]

where \(n = 2\) for null geodesics and \(n = 3\) for time–like geodesics. Here \(v\) is an affine parameter along a congruence of geodesics, with tangent vector \(l^\alpha\) which are hypersurface orthogonal. The quantity \(\rho\) is the average rate of convergence of the geodesics, while \(\sigma\) measures the shear. The term \(R_{\alpha\beta} l^\alpha l^\beta\) gives the direct gravitational effect of the matter on the convergence of the geodesics. By the Einstein equation, \(G_{\alpha\beta} = 8\pi T_{\alpha\beta}\) (here \(G_{\alpha\beta}\) is the Einstein tensor), the term \(R_{\alpha\beta} l^\alpha l^\beta\) will be non–negative for any null vector \(l^\alpha\) if the matter obeys the so–called weak energy condition, which says that the energy density \(T_{00}\) is non–negative in any frame, i.e.,

\[
T_{\alpha\beta} v^\alpha v^\beta \geq 0,
\]

for any time–like vector \(v^\alpha\), is obeyed by the classical SEM–tensor of any reasonable matter [Hawking and Israel (1979) Hawking and Penrose (1996)].

Suppose the weak energy condition holds, and that the null geodesics from a point \(p\) begin to converge again and that \(\rho\) has the positive value \(\rho_0\). Then the Newman–Penrose equation would imply that the convergence \(\rho\) would become infinite at a point \(q\) within an affine parameter distance \(\frac{1}{\rho_0}\) if the null geodesic can be extended that far. If \(\rho = \rho_0\) at \(v = v_0\) then \(\rho \geq \frac{1}{\rho_0} + v_0 - v\). Thus there is a conjugate point before \(v = v_0 + \rho^{-1}\).

Infinitesimally neighboring null geodesics from \(p\) will intersect at \(q\). This means the point \(q\) will be conjugate to \(p\) along the null geodesic \(\gamma\) joining them. For points on \(\gamma\) beyond the conjugate point \(q\) there will be a variation of \(\gamma\) that gives a time like curve from \(p\). Thus \(\gamma\) cannot lie in the boundary of the future of \(p\) beyond the conjugate point \(q\). So \(\gamma\) will have a future end point as a generator of the boundary of the future of \(p\).

The situation with time–like geodesics is similar, except that the strong

---

\(^6\)Recall that the Newman–Penrose formalism is a set of notation developed by E.T. Newman and R. Penrose [Newman and Penrose (1962) Teukolsky (1973)] for general relativity (GR). Their notation is an effort to treat GR in terms of spinor notation, which introduces complex forms of the usual variables used in GR. The most often–used variables in their formalism are the Weyl scalars, derived from the Weyl tensor.
energy condition \cite{Hawking and Israel (1979), Hawking and Penrose (1996)},

\[ T_{\alpha\beta} v^\alpha v^\beta \geq \frac{1}{2} v^\alpha v^\beta T, \]  

that is required to make \( R_{\alpha\beta} l^\alpha l^\beta \) non-negative for every time like vector \( l^\alpha \), is rather stronger than the weak energy condition (4.58). However, it is still physically reasonable, at least in an averaged sense, in classical theory. If the strong energy condition holds, and the time like geodesics from \( p \) begin converging again, then there will be a point \( q \) conjugate to \( p \).

Finally there is the generic energy condition, which says:

1. The strong energy condition holds.
2. Every time–like or null geodesic has a point where \( l_a R_{bcdf} l^c l^d \neq 0 \).

One normally thinks of a space–time singularity as a region in which the curvature becomes unboundedly large. However, the trouble with this definition is that one could simply leave out the singular points and say that the remaining manifold was the whole of space–time. It is therefore better to define space–time as the maximal manifold on which the metric is suitably smooth. One can then recognize the occurrence of singularities by the existence of incomplete geodesics that cannot be extended to infinite values of the affine parameter.

**Hawking–Penrose Singularity Theorems**

Hawking–Penrose Singularity is defined as follows \cite{Hawking and Israel (1979), Penrose (1967), Hawking and Penrose (1996)}:

A space–time manifold is singular if it is time–like or null geodesically incomplete but cannot be embedded in a larger space–time manifold.

This definition reflects the most objectionable feature of singularities, that there can be particles whose history has a beginning or end at a finite time. There are examples in which geodesic incompleteness can occur with the curvature remaining bounded, but it is thought that generically the curvature will diverge along incomplete geodesics. This is important if one is to appeal to quantum effects to solve the problems raised by singularities in classical general relativity.

**Singularity Theorems include:**

1. Energy condition (i.e., weak \( 4.58 \), strong \( 4.59 \), or generic \( 4.3.3 \)).
2. Condition on global structure (e.g., there should not be any closed time–like curves).
Gravity strong enough to trap a region (so that nothing could escape).

The various singularity theorems show that space–time must be time like or null geodesically incomplete if different combinations of the three kinds of conditions hold. One can weaken one condition if one assumes stronger versions of the other two. The Hawking–Penrose Singularity theorems have the generic energy condition, the strongest of the three energy conditions. The global condition is fairly weak, that there should be no closed time like curves. And the no escape condition is the most general, that there should be either a trapped surface or a closed space like three surface.

The theorems predict singularities in two situations. One is in the future in the gravitational collapse of stars and other massive bodies. Such singularities would be an end of time, at least for particles moving on the incomplete geodesics. The other situation in which singularities are predicted is in the past at the beginning of the present expansion of the universe.

The prediction of singularities means that classical general relativity is not a complete theory. Because the singular points have to be cut out of the space–time manifold one cannot define the field equations there and cannot predict what will come out of a singularity. With the singularity in the past the only way to deal with this problem seems to be to appeal to quantum gravity. But the singularities that are predicted in the future seem to have a property that Penrose has called, Cosmic Censorship. That is they conveniently occur in places like black holes that are hidden from external observers. So any break down of predictability that may occur at these singularities will not affect what happens in the outside world, at least not according to classical theory.

Hawking Cosmic Censorship Hypothesis says: “Nature abhors a naked singularity” [Hawking and Israel (1979); Hawking and Penrose (1996)]. However, there is unpredictability in the quantum theory. This is related to the fact that gravitational fields can have intrinsic entropy which is not just the result of coarse graining. Gravitational entropy, and the fact that time has a beginning and may have an end, are the two main themes of Hawking’s research, because they are the ways in which gravity is distinctly different from other physical fields.

The fact that gravity has a quantity that behaves like entropy was first noticed in the purely classical theory. It depends on Penrose’s Cosmic Censorship Conjecture. This is unproved but is believed to be true for suitably general initial data and state equations.

One makes the approximation of treating the region around a collapsing
star as asymptotically flat. Then, as Penrose showed, one can conformally
embed the space–time manifold $M$ in a manifold with boundary $\bar{M}$. The
boundary $\partial M$ will be a null surface and will consist of two components,
future and past null infinity, called $I^+$ and $I^-$. One says that weak Cosmic
Censorship holds if two conditions are satisfied. First, it is assumed that the
null geodesic generators of $I^+$ are complete in a certain conformal metric.
This implies that observers far from the collapse live to an old age and are
not wiped out by a thunderbolt singularity sent out from the collapsing
star. Second, it is assumed that the past of $I^+$ is globally hyperbolic.
This means there are no naked singularities that can be seen from large
distances. Penrose has also a stronger form of Cosmic Censorship which
assumes that the whole space–time is globally hyperbolic.

Weak Cosmic Censorship Hypothesis reads:

1. $I^+$ and $I^-$ are complete.
2. $I^-(I^+)$ is globally hyperbolic.

If weak Cosmic Censorship holds, the singularities that are predicted
to occur in gravitational collapse cannot be visible from $I^+$. This means
that there must be a region of space–time that is not in the past of $I^+$.
This region is said to be a black hole because no light or anything else can
escape from it to infinity. The boundary of the black hole region is called the
event horizon. Because it is also the boundary of the past of $I^+$ the
event horizon will be generated by null–geodesic segments that may have
past end points but don’t have any future end points. It then follows that
if the weak energy condition holds the generators of the horizon cannot be
converging. For if they were they would intersect each other within a finite
distance [Hawking and Israel (1979)] [Penrose (1967)] [Hawking and Penrose
(1996)].

This implies that the area of a cross section of the event horizon can
never decrease with time and in general will increase. Moreover if two
black holes collide and merge together the area of the final black hole will
be greater than the sum of the areas of the original black holes. This is
very similar to the behavior of entropy according to the Second Law of
Thermodynamics:

$$\delta A \geq 0.$$  

Second Law of Black Hole Mechanics: $\delta A \geq 0$.

Second Law of Thermodynamics: $\delta S \geq 0$.

The similarity with thermodynamics is increased by what is called the

---

Recall that Second Law of Thermodynamics states: Entropy can never decrease and
the entropy of a total system is greater than the sum of its constituent parts.
First Law of Black Hole Mechanics, which relates the change in mass of a black hole to the change in the area of the event horizon and the change in its angular momentum and electric charge. One can compare this to the First Law of Thermodynamics which gives the change in internal energy in terms of the change in entropy and the external work done on the system. [Hawking and Israel (1979)] [Hawking and Penrose (1996)]

First Law of Black Hole Mechanics: \( \delta E = \frac{\kappa}{8\pi} \delta A + \Omega \delta J + \Phi \delta Q. \)

First Law of Thermodynamics: \( \delta E = T \delta S + P \delta V. \)

One sees that if the area \( A \) of the event horizon is analogous to entropy \( S \) then the quantity analogous to temperature is what is called the surface gravity of the black hole \( \kappa \). This is a measure of the strength of the gravitational field on the event horizon. The similarity with thermodynamics is further increased by the so-called Zeroth Law of Black Hole Mechanics: the surface gravity is the same everywhere on the event horizon of a time independent black hole [Hawking and Israel (1979)].

Zeroth Law of Black Hole Mechanics: \( \kappa \) is the same everywhere on the horizon of a time independent black hole.

Zeroth Law of Thermodynamics: \( T \) is the same everywhere for a system in thermal equilibrium.

Encouraged by these similarities Bekenstein proposed that some multiple of the area of the event horizon actually was the entropy of a black hole. He suggested a generalized Second Law: the sum of this black hole entropy and the entropy of matter outside black holes would never decrease (see [Strominger and Vafa (1996)])

Generalized Second Law: \( \delta (S + cA) \geq 0. \)

However, this proposal was not consistent. If black holes have an entropy proportional to horizon area \( A \) they should also have a non zero temperature proportional to surface gravity.

Path–Integral Model for Black Holes

Recall that the fact that gravity is attractive means that it will tend to draw the matter in the universe together to form objects like stars and galaxies. These can support themselves for a time against further contraction by thermal pressure, in the case of stars, or by rotation and internal motions, in the case of galaxies. However, eventually the heat or the angular momentum will be carried away and the object will begin to shrink. If the mass is less than about one and a half times that of the Sun the contraction can be stopped by the degeneracy pressure of electrons or neutrons. The object
will settle down to be a white dwarf or a neutron star respectively. However, if the mass is greater than this limit there is nothing that can hold it up and stop it continuing to contract. Once it has shrunk to a certain critical size the gravitational field at its surface will be so strong that the light cones will be bent inward [Hawking and Israel (1979); Hawking and Penrose (1996)].

If the Cosmic Censorship Conjecture is correct the trapped surface and the singularity it predicts cannot be visible from far away. Thus there must be a region of space–time from which it is not possible to escape to infinity. This region is said to be a black hole. Its boundary is called the event horizon and it is a null surface formed by the light rays that just fail to get away to infinity. As we saw in the last subsection, the area $A$ of a cross section of the event horizon can never decrease, at least in the classical theory. This, and perturbation calculations of spherical collapse, suggest that black holes will settle down to a stationary state.

Recall that the Schwarzschild metric form, given by

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \left(1 - \frac{2M}{r}\right)^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2),$$

represents the gravitational field that a black hole would settle down to if it were non rotating. In the usual $r$ and $t$ coordinates there is an apparent singularity at the Schwarzschild radius $r = 2M$. However, this is just caused by a bad choice of coordinates. One can choose other coordinates in which the metric is regular there.

Now, if one performs the Wick rotation, $t = i\tau$, one gets a positive definite metric, usually called Euclidean even though they may be curved. In the Euclidean–Schwarzschild metric

$$ds^2 = x^2 \left(\frac{d\tau}{4M}\right)^2 + \left(\frac{\nu^2}{4M^2}\right)^2 dx^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2),$$

there is again an apparent singularity at $r = 2M$. However, one can define a new radial coordinate $x$ to be $4M(1 - 2Mr^{-1})^{1/2}$.

The metric in the $x - \tau$ plane then becomes like the origin of polar coordinates if one identifies the coordinate $\tau$ with period $8\pi M$. Similarly, other Euclidean black hole metrics will have apparent singularities on their horizons which can be removed by identifying the imaginary time coordinate with period $\frac{2\pi}{\beta}$.

To see the significance of having imaginary time identified with some period $\beta$, let us consider the amplitude to go from some field configuration $\phi_1$ on the surface $t_1$ to a configuration $\phi_2$ on the surface $t_2$. This will be
given by the matrix element of $e^{iH(t_2-t_1)}$. However, one can also represent this amplitude as a path integral over all fields $\phi$ between $t_1$ and $t_2$ which agree with the given fields $\phi_1$ and $\phi_2$ on the two surfaces,

$$\langle \phi_2, t_2 | \phi_1, t_1 \rangle = \langle \phi_2 | \exp(-iH(t_2 - t_1)) | \phi_1 \rangle = \int D[\phi] \exp(iA[\phi]).$$

One now chooses the time separation $(t_2 - t_1)$ to be pure imaginary and equal to $\beta$. One also puts the initial field $\phi_1$ equal to the final field $\phi_2$ and sums over a complete basis of states $\phi_n$. On the left one has the expectation value of $e^{-\beta H}$ summed over all states. This is just the thermodynamic partition function $Z$ at the temperature $T = \beta^{-1},$

$$\langle \phi_2 | \phi_1 \rangle = \sum_n \langle \phi_n | \exp(-\beta H) | \phi_n \rangle = \int D[\phi] \exp(-A[\phi]). \quad (4.60)$$

On the r.h.s. of this equation one has a path integral (see chapter 6). One puts $\phi_1 = \phi_2$ and sums over all field configurations $\phi_n$. This means that effectively one is doing the path integral over all fields $\phi$ on a space–time that is identified periodically in the imaginary time direction with period $\beta$. Thus the partition function for the field $\phi$ at temperature $T$ is given by a path integral over all fields on a Euclidean space–time. This space–time is periodic in the imaginary time direction with period $\beta = T^{-1}$.

If one calculates the path integral in flat space–time identified with period $\beta$ in the imaginary time direction one gets the usual result for the partition function of black body radiation. However, as we have just seen, the Euclidean–Schwarzschild solution is also periodic in imaginary time with period $\frac{2\pi}{\kappa}$. This means that fields on the Schwarzschild background will behave as if they were in a thermal state with temperature $\frac{\kappa}{2\pi}$.

The periodicity in imaginary time explained why the messy calculation of frequency mixing led to radiation that was exactly thermal. However, this derivation avoided the problem of the very high frequencies that take part in the frequency mixing approach. It can also be applied when there are interactions between the quantum fields on the background. The fact that the path integral is on a periodic background implies that all physical quantities like expectation values will be thermal. This would have been very difficult to establish in the frequency mixing approach [Hawking and Israel (1979); Hawking and Penrose (1996)].

One can extend these interactions to include interactions with the gravitational field itself. One starts with a background metric $g_0$ such as the
Euclidean–Schwarzschild metric that is a solution of the classical field equations. One can then expand the action $A$ in a power series in the perturbations $\delta g$ about $g_0$, as

$$A[g] = A[g_0] + A_2(\delta g)^2 + A_3(\delta g)^3 + ...$$

Here, the linear term vanishes because the background is a solution of the field equations. The quadratic term can be regarded as describing gravitons on the background while the cubic and higher terms describe interactions between the gravitons. The path integral over the quadratic terms are finite. There are non renormalizable divergences at two loops in pure gravity but these cancel with the fermions in super–gravity theories. It is not known whether super–gravity theories have divergences at three loops or higher because no one has been brave or foolhardy enough to try the calculation. Some recent work indicates that they may be finite to all orders. But even if there are higher loop divergences they will make very little difference except when the background is curved on the scale of the Planck length ($10^{-33}$ cm).

More interesting than the higher order terms is the zeroth order term, the action of the background metric $g_0$ [Hawking and Israel (1979); Hawking and Penrose (1996)],

$$A = -\frac{1}{16\pi} \int R(-g)^{\frac{1}{2}} \, d^4x + \frac{1}{8\pi} \int K(\pm h)^{\frac{1}{2}} \, d^3x.$$  

Recall that the usual Einstein–Hilbert action for general relativity is the volume integral of the scalar curvature $R$. This is zero for vacuum solutions so one might think that the action of the Euclidean-Schwarzschild solution was zero. However, there is also a surface term in the action proportional to the integral of $K$, the trace of the second fundamental form of the boundary surface. When one includes this and subtracts off the surface term for flat space one finds the action of the Euclidean–Schwarzschild metric is $-\frac{\beta^2}{16\pi}$ where $\beta$ is the period in imaginary time at infinity. Thus the dominant contribution to the path integral for the partition function $Z$ given by (4.60), is $e^{\frac{\beta^2}{16\pi}}$,

$$Z = \sum \exp(-\beta E_n) = \exp \left( -\frac{\beta^2}{16\pi} \right).$$

If one differentiates $\log Z$ with respect to the period $\beta$ one gets the
expectation value of the energy, or in other words, the mass,

\[ \langle E \rangle = -\frac{d}{d\beta} (\log Z) = \frac{\beta}{8\pi}. \]

So this gives the mass \( M = \frac{\beta}{8\pi} \). This confirms the relation between the mass and the period, or inverse temperature, that we already knew. However, one can go further. By standard thermodynamic arguments, the log of the partition function is equal to minus the free energy \( F \) divided by the temperature \( T \), i.e., \( \log Z = -\frac{F}{T} \). And the free energy is the mass or energy plus the temperature times the entropy \( S \), i.e., \( F = \langle E \rangle + TS \). Putting all this together one sees that the action of the black hole gives an entropy of \( 4\pi M^2 \),

\[ S = \frac{\beta^2}{16\pi} = 4\pi M^2 = \frac{1}{4 \mathcal{A}}. \]

This is exactly what is required to make the laws of black holes the same as the laws of thermodynamics [Hawking and Israel (1979); Hawking and Penrose (1996)]. The reason why does one get this intrinsic gravitational entropy which has no parallel in other quantum field theories, is that gravity allows different topologies for the space–time manifold.

In the case we are considering the Euclidean–Schwarzschild solution has a boundary at infinity that has topology \( S^2 \times S^1 \). The \( S^2 \) is a large space like two sphere at infinity and the \( S^1 \) corresponds to the imaginary time direction which is identified periodical. One can fill in this boundary with metrics of at least two different topologies. One is the Euclidean–Schwarzschild metric. This has topology \( R^2 \times S^2 \), that is the Euclidean two plane times a two sphere. The other is \( R^3 \times S^1 \), the topology of Euclidean flat space periodically identified in the imaginary time direction. These two topologies have different Euler numbers. The Euler number of periodically identified flat space is zero, while that of the Euclidean–Schwarzschild solution is two,

\[ \text{Total action} = M(\tau_2 - \tau_1). \]

The significance of this is as follows: on the topology of periodically identified flat space one can find a periodic time function \( \tau \) whose gradient is no where zero and which agrees with the imaginary time coordinate on the boundary at infinity. One can then work out the action of the region between two surfaces \( \tau_1 \) and \( \tau_2 \). There will be two contributions to the
action, a volume integral over the matter Lagrangian, plus the Einstein–Hilbert Lagrangian and a surface term. If the solution is time independent the surface term over \( \tau = \tau_1 \) will cancel with the surface term over \( \tau = \tau_2 \). Thus the only net contribution to the surface term comes from the boundary at infinity. This gives half the mass times the imaginary time interval \((\tau_2 - \tau_1)\). If the mass is non–zero there must be non–zero matter fields to create the mass. One can show that the volume integral over the matter Lagrangian plus the Einstein–Hilbert Lagrangian also gives \( \frac{1}{2} M (\tau_2 - \tau_1) \). Thus the total action is \( M (\tau_2 - \tau_1) \). If one puts this contribution to the log of the partition function into the thermodynamic formulae one finds the expectation value of the energy to be the mass, as one would expect. However, the entropy contributed by the background field will be zero.

However, the situation is different with the Euclidean–Schwarzschild solution, which says:

\[
\text{Total action including corner contribution} = M (\tau_2 - \tau_1).
\]

\[
\text{Total action without corner contribution} = \frac{1}{2} M (\tau_2 - \tau_1).
\]

Because the Euler number is two rather than zero one cannot find a time function \( \tau \) whose gradient is everywhere non–zero. The best one can do is choose the imaginary time coordinate of the Schwarzschild solution. This has a fixed two sphere at the horizon where \( \tau \) behaves like an angular coordinate. If one now works out the action between two surfaces of constant \( \tau \) the volume integral vanishes because there are no matter fields and the scalar curvature is zero. The trace \( K \) surface term at infinity again gives \( \frac{1}{2} M (\tau_2 - \tau_1) \). However there is now another surface term at the horizon where the \( \tau_1 \) and \( \tau_2 \) surfaces meet in a corner. One can evaluate this surface term and find that it also is equal to \( \frac{1}{2} M (\tau_2 - \tau_1) \). Thus the total action for the region between \( \tau_1 \) and \( \tau_2 \) is \( M (\tau_2 - \tau_1) \). If one used this action with \( \tau_2 - \tau_1 = \beta \) one would find that the entropy was zero. However, when one looks at the action of the Euclidean Schwarzschild solution from a 4–dimensional point of view rather than a 3 + 1, there is no reason to include a surface term on the horizon because the metric is regular there. Leaving out the surface term on the horizon reduces the action by one quarter the area of the horizon, which is just the intrinsic gravitational entropy of the black hole \cite{Hawking:1979}. 

\cite{Hawking:1979}, \cite{Hawking:1996}. 


Quantum Cosmology

According to Hawking, cosmology used to be considered a pseudo-science and the preserve of physicists who may have done useful work in their earlier years but who had gone mystic in their dotage. There is a serious objection that cosmology cannot predict anything about the universe unless it makes some assumption about the initial conditions. Without such an assumption, all one can say is that things are as they are now because they were as they were at an earlier stage. Yet many people believe that science should be concerned only with the local laws which govern how the universe evolves in time. They would feel that the boundary conditions for the universe that determine how the universe began were a question for metaphysics or religion rather than science [Hawking and Israel (1979); Hawking and Penrose (1996)].

Hawking–Penrose theorems showed that according to general relativity there should be a singularity in our past. At this singularity the field equations could not be defined. Thus classical general relativity brings about its own downfall: it predicts that it cannot predict the universe. For Hawking this sounds rally disturbing: If the laws of physics could break down at the beginning of the universe, why couldn’t they break down any where. In quantum theory it is a principle that anything can happen if it is not absolutely forbidden. Once one allows that singular histories could take part in the path integral they could occur any where and predictability would disappear completely. If the laws of physics break down at singularities, they could break down any where.

The only way to have a scientific theory is if the laws of physics hold everywhere including at the beginning of the universe. One can regard this as a triumph for the Principle of Democracy: Why should the beginning of the universe be exempt from the laws that apply to other points. If all points are equal one cannot allow some to be more equal than others.

To implement the idea that the laws of physics hold everywhere, one should take the path integral only over non–singular metrics. One knows in the ordinary path integral case that the measure is concentrated on non–differentiable paths. But these are the completion in some suitable topology of the set of smooth paths with well defined action. Similarly, one would expect that the path integral for quantum gravity should be taken over the completion of the space of smooth metrics. What the path integral cannot include is metrics with singularities whose action is not defined.

In the case of black holes we saw that the path integral should be taken
over Euclidean, that is, positive definite metrics. This meant that the singularities of black holes, like the Schwarzschild solution, did not appear on the Euclidean metrics which did not go inside the horizon. Instead the horizon was like the origin of polar coordinates. The action of the Euclidean metric was therefore well defined. One could regard this as a quantum version of Cosmic Censorship: the break down of the structure at a singularity should not affect any physical measurement.

It seems, therefore, that the path integral for quantum gravity should be taken over non–singular Euclidean metrics. But what should the boundary conditions be on these metrics. There are two, and only two, natural choices. The first is metrics that approach the flat Euclidean metric outside a compact set. The second possibility is metrics on manifolds that are compact and without boundary. Therefore, the natural choices for path integral for quantum gravity are [Hawking and Israel (1979); Hawking and Penrose (1996)]: (i) asymptotically Euclidean metrics, and (ii) compact metrics without boundary. The first class of asymptotically Euclidean metrics is appropriate for scattering calculations. In these one sends particles in from infinity and observes what comes out again to infinity. All measurements are made at infinity where one has a flat background metric and one can interpret small fluctuations in the fields as particles in the usual way. One doesn’t ask what happens in the interaction region in the middle. That is why one does a path integral over all possible histories for the interaction region, that is, over all asymptotically Euclidean metrics. However, in cosmology one is interested in measurements that are made in a finite region rather than at infinity. We are on the inside of the universe not looking in from the outside. To see what difference this makes let us first suppose that the path integral for cosmology is to be taken over all asymptotically Euclidean metrics.

The so–called No Boundary Proposal of Hartle and Hawking reads [Hawking and Israel (1979); Hawking and Penrose (1996)]: The path integral for quantum gravity should be taken over all compact Euclidean metrics. One can paraphrase this as: the boundary condition of the universe is that it has no boundary. According to Hawking, this no boundary proposal seems to account for the universe we live in. That is an isotropic and homogeneous expanding universe with small perturbations. We can observe the spectrum and statistics of these perturbations in the fluctuations in the microwave background. The results so far agree with the predictions of the no boundary proposal. It will be a real test of the proposal and the whole Euclidean quantum gravity program when the observations of the
microwave background are extended to smaller angular scales.

In order to use the no boundary proposal to make predictions, it is useful to introduce a concept that can describe the state of the universe at one time:

\[
\text{Probability of induced metric } h_{ij} \text{ on } \Sigma = \int_{\text{metrics on } M \text{ that induce } h_{ij} \text{ on } \Sigma} d[g] \exp(-A[g]).
\]

Consider the probability that the space–time manifold \( M \) contains an embedded three dimensional manifold \( \Sigma \) with induced metric \( h_{ij} \). This is given by a path integral over all metrics \( g_{ab} \) on \( M \) that induce \( h_{ij} \) on \( \Sigma \). If \( M \) is simply–connected, which we will assume, the surface \( \Sigma \) will divide \( M \) into two parts \( M^+ \) and \( M^- \). \[\text{Hawking and Israel (1979); Hawking and Penrose (1996)}\] 

\[
\text{Probability of } h_{ij} = \Psi^+(h_{ij}) \times \Psi^-(h_{ij}), \quad \text{where} \quad 
\Psi^+(h_{ij}) = \int_{\text{metrics on } M^\star \text{ that induce } h_{ij} \text{ on } \Sigma} d[g] \exp(-A[g]).
\]

In this case, the probability for \( \Sigma \) to have the metric \( h_{ij} \) can be factorized. It is the product of two wave functions \( \Psi^+ \) and \( \Psi^- \). These are given by path integrals over all metrics on \( M^+ \) and \( M^- \) respectively, that induce the given three metric \( h_{ij} \) on \( \Sigma \). In most cases, the two wave functions will be equal and we will drop the superscripts \(+\) and \(-\). \( \Psi \) is called the wave function of the universe. If there are matter fields \( \phi \), the wave function will also depend on their values \( \phi_0 \) on \( \Sigma \). But it will not depend explicitly on time because there is no preferred time coordinate in a closed universe. The no boundary proposal implies that the wave function of the universe is given by a path integral over fields on a compact manifold \( M^+ \) whose only boundary is the surface \( \Sigma \). The path integral is taken over all metrics and matter fields on \( M^+ \) that agree with the metric \( h_{ij} \) and matter fields \( \phi_0 \) on \( \Sigma \).

One can describe the position of the surface \( \Sigma \) by a function \( \tau \) of three coordinates \( x_i \) on \( \Sigma \). But the wave function defined by the path integral cannot depend on \( \tau \) or on the choice of the coordinates \( x_i \). This implies that the wave function \( \Psi \) has to obey four functional differential equations. Three of these equations are called the momentum constraint. One can describe the position of the surface \( \Sigma \) by a function \( \tau \) of three coordinates \( x_i \) on \( \Sigma \). But the wave function defined by the path integral cannot depend on \( \tau \) or on the choice of the coordinates \( x_i \). This implies that the wave
function $\Psi$ has to obey four functional differential equations. Three of these equations are called the momentum constraint equation:

$$\left( \frac{\partial g}{\partial h_{ij}} \right)_{,j} = 0.$$  

They express the fact that the wave function should be the same for different 3 metrics $h_{ij}$ that can be get from each other by transformations of the coordinates $x_i$. The fourth equation is called the Wheeler–DeWitt equation

$$\left( G_{ijkl} \frac{\partial^2}{\partial h_{ij}\partial h_{kl}} - h^{\frac{3}{2}} R \right) \Psi = 0.$$  

It corresponds to the independence of the wave function on $\tau$. One can think of it as the Schrödinger equation for the universe. But there is no time derivative term because the wave function does not depend on time explicitly.

In order to estimate the wave function of the universe, one can use the saddle point approximation to the path integral as in the case of black holes. One finds a Euclidean metric $g_0$ on the manifold $M^+$ that satisfies the field equations and induces the metric $h_{ij}$ on the boundary $\Sigma$. One can then expand the action $A$ in a power series around the background metric $g_0$,

$$A[g] = A[g_0] + \frac{1}{2} \delta g A_2 \delta g + ...$$

As before, the term linear in the perturbations vanishes. The quadratic term can be regarded as giving the contribution of gravitons on the background and the higher order terms as interactions between the gravitons. These can be ignored when the radius of curvature of the background is large compared to the Planck scale. Therefore, according to [Hawking and Israel (1979); Hawking and Penrose (1996)] we have

$$\Psi \approx \frac{1}{(\det A_2)^{\frac{1}{2}}} \exp(-A[g_0]).$$

Consider now a situation in which there are no matter fields but there is a positive cosmological constant $\Lambda$. Let us take the surface $\Sigma$ to be a three sphere and the metric $h_{ij}$ to be the round three sphere metric of radius $a$. Then the manifold $M^+$ bounded by $\Sigma$ can be taken to be the four ball. The metric that satisfies the field equations is part of a four sphere of radius $1/\pi$ where $H^2 = \frac{\Lambda}{3}$,

$$A = \frac{1}{16\pi} \int (R - 2\Lambda)(-g)^{\frac{1}{2}} d^4x + \frac{1}{8\pi} \int K(\pm h)^{\frac{1}{2}} d^3x.$$
For a 3-sphere $\Sigma$ of radius less than $\frac{1}{H}$ there are two possible Euclidean solutions: either $M^+$ can be less than a hemisphere or it can be more. However there are arguments that show that one should pick the solution corresponding to less than a hemisphere.

One can interpret the wave function $\Psi$ as follows. The real time solution of the Einstein equations with a $\Lambda$ term and maximal symmetry is de Sitter space (see, e.g., [Witten (1998b)]). This can be embedded as a hyperboloid in five dimensional Minkowski space. Here, we have two choices:

(1) **Lorentzian–de Sitter metric**, 
\[
\begin{align*}
    ds^2 &= -dt^2 + \frac{1}{H^2} \cosh Ht (dr^2 + \sin^2 r (d\theta^2 + \sin^2 \theta d\phi^2)).
\end{align*}
\]

One can think of it as a closed universe that shrinks down from infinite size to a minimum radius and then expands again exponentially. The metric can be written in the form of a Friedmann universe with scale factor $\cosh Ht$. Putting $\tau = it$ converts the $\cosh$ into $\cos$ giving the Euclidean metric on a four sphere of radius $\frac{1}{H}$.

(2) **Euclidean metric**, 
\[
\begin{align*}
    ds^2 &= d\tau^2 + \frac{1}{H^2} \cos H\tau (dr^2 + \sin^2 r (d\theta^2 + \sin^2 \theta d\phi^2)).
\end{align*}
\]

Thus one gets the idea that a wave function which varies exponentially with the three metric $h_{ij}$ corresponds to an imaginary time Euclidean metric. On the other hand, a wave function which oscillates rapidly corresponds to a real time Lorentzian metric.

Hawking says: “The Euclidean path integral over all topologically trivial metrics can be done by time slicing and so is unitary when analytically continued to the Lorentzian. On the other hand, the path integral over all topologically non–trivial metrics is asymptotically independent of the initial state. Thus the total path integral is unitary and information is not lost in the formation and evaporation of black holes. The way the information gets out seems to be that a true event horizon never forms, just an apparent horizon.”

Like in the case of the pair creation of black holes, one can describe the spontaneous creation of an exponentially expanding universe. One joins the lower half of the Euclidean four sphere to the upper half of the Lorentzian hyperboloid.
Unlike the black hole pair creation, one couldn’t say that the de Sitter universe was created out of field energy in a pre–existing space. Instead, it would quite literally be created out of nothing: not just out of the vacuum but out of absolutely nothing at all because there is nothing outside the universe. In the Euclidean regime, the de Sitter universe is just a closed space like the surface of the Earth but with two more dimensions (Witten (1998b)). If the cosmological constant is small compared to the Planck value, the curvature of the Euclidean four sphere should be small. This will mean that the saddle point approximation to the path integral should be good, and that the calculation of the wave function of the universe will not be affected by our ignorance of what happens in very high curvatures.

One can also solve the field equations for boundary metrics that are not exactly the round three sphere metric. If the radius of the three sphere is less than \( \frac{1}{H} \), the solution is a real Euclidean metric. The action will be real and the wave function will be exponentially damped compared to the round three sphere of the same volume. If the radius of the three sphere is greater than this critical radius there will be two complex conjugate solutions and the wave function will oscillate rapidly with small changes in \( h_{ij} \).

Any measurement made in cosmology can be formulated in terms of the wave function. Thus the no boundary proposal makes cosmology into a science because one can predict the result of any observation. The case we have just been considering of no matter fields and just a cosmological constant does not correspond to the universe we live in. Nevertheless, it is a useful example, both because it is a simple model that can be solved fairly explicitly and because, as we shall see, it seems to correspond to the early stages of the universe.

Although it is not obvious from the wave function, a de Sitter universe has thermal properties rather like a black hole. One can see this by writing the de Sitter metric in a static form (rather like the Schwarzschild solution)

\[
ds^2 = -(1 - H^2 r^2) dt^2 + (1 - H^2 r^2)^{-1} dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\phi^2).\]

There is an apparent singularity at \( r = \frac{1}{H} \). However, as in the Schwarzschild solution, one can remove it by a coordinate transformation and it corresponds to an event horizon.

If one returns to the static form of the de Sitter metric and put \( \tau = it \) one gets a Euclidean metric. There is an apparent singularity on the horizon. However, by defining a new radial coordinate and identifying \( \tau \) with period \( \frac{2\pi}{H} \), one gets a regular Euclidean metric which is just the four sphere. Because the imaginary time coordinate is periodic, de Sit-
ter space and all quantum fields in it will behave as if they were at a temperature $\frac{H}{2\pi}$. As we shall see, we can observe the consequences of this temperature in the fluctuations in the microwave background. One can also apply arguments similar to the black hole case to the action of the Euclidean–de Sitter solution [Witten (1998b)]. One finds that it has an intrinsic entropy of $\frac{\pi}{H}$, which is a quarter of the area of the event horizon. Again this entropy arises for a topological reason: the Euler number of the four sphere is two. This means that there cannot be a global time coordinate on Euclidean–de Sitter space. One can interpret this cosmological entropy as reflecting an observer’s lack of knowledge of the universe beyond his event horizon [Hawking and Israel (1979); Hawking and Penrose (1996)].

4.3.4 Generalized Quantum Mechanics

Recall that familiar textbook quantum mechanics assumes a fixed background space–time to define states on space–like surfaces and their unitary evolution between them. Quantum theory has changed as our conceptions of space and time have evolved. But quantum mechanics needs to be generalized further for quantum gravity where space–time geometry is fluctuating and without definite value. In this section, following [Hartle (2005b)], we review a fully 4D, sum–over–histories, generalized quantum mechanics of cosmological space–time geometry. This generalization is constructed within the framework of generalized quantum theory. This is a minimal set of principles for quantum theory abstracted from the modern quantum mechanics of closed systems, most generally the universe. In this generalization, states of fields on space–like surfaces and their unitary evolution are emergent properties appropriate when space–time geometry behaves approximately classically. The principles of generalized quantum theory allow for the further generalization that would be necessary were space–time not fundamental. Emergent space–time phenomena are discussed in general and illustrated with the example of the classical space–time geometries with large space–like surfaces that emerge from the no–boundary wave function of the universe.

Does quantum mechanics apply to space–time? This is an old question.
Belgian physicist L. Rosenfeld wrote one of the first papers on quantum gravity [Rosenfeld (1930)], but late in his career came to the conclusion that the quantization of the gravitational field would be meaningless [Rosenfeld (1963); Rosenfeld (1966)]. Today, there are probably more colleagues of the opinion that quantum theory needs to be replaced than there are who think that it doesn’t apply to space–time. But in the end this is an experimental question as Rosenfeld stressed.

The answer that J. Hartle proposes is: “Quantum mechanics can be applied to space–time provided that the usual textbook formulation of quantum theory is suitably generalized.” A generalization is necessary because, in one way or another, the usual formulations rely on a fixed space–time geometry to define states on space–like surfaces and the time in which they evolve unitarily one surface to another. But in a quantum theory of gravity, space–time geometry is generally fluctuating and without definite value. The usual formulations are emergent from a more general perspective when geometry is approximately classical and can supply the requisite fixed notions of space and time.

A framework for investigating generalizations of usual quantum mechanics can be abstracted from the modern quantum mechanics of closed systems [Griffiths (2002); Omnès (1994); Gell-Mann (1994)] which enables quantum mechanics to be applied to cosmology. The resulting framework, the so–called generalized quantum theory [Hartle (1991a); Hartle (1995b); Isham (1994)], defines a broad class of generalizations of usual quantum mechanics.

A generalized quantum theory of a physical system (most generally the universe) is built on three elements which can be very crudely characterized as follows [Hartle (2005b)]:

- The possible fine–grained descriptions of the system.
- The coarse–grained descriptions constructed from the fine–grained ones.
- A measure of the quantum interference between different coarse–grained descriptions incorporating the principle of superposition

The standard quantum two–slit experiment provides an immediate, concrete illustration. A set of possible fine–grained descriptions of an electron moving through the two–slit apparatus are its Feynman paths in time.

---

8Rosenfeld considered the example of classical geometry curved by the expected value of the stress–energy of quantum fields. Some of the difficulties with this proposal, including experimental inconsistencies, are discussed in [Page and Geilker (1981)].
Quantum Leap

(histories) from the source to the detecting screen. One coarse-grained description is by which slit the electron went through on its way to detection in an interval $\Delta$ about a position $y$ on the screen at a later time. Amplitudes $\psi_U(y)$ and $\psi_L(y)$ for the two coarse-grained histories where the electron goes through the upper or lower slit and arrives at a point $y$ on the screen can be computed as a sum over paths in the usual way. The natural measure of interference between these two histories is the overlap of these two amplitudes integrated over the interval $\Delta$ in which the electron is detected. In this way usual quantum mechanics is a special case of generalized quantum theory.

Probabilities cannot be assigned to the two coarse-grained histories in the two-slit experiment because they interfere. The probability to arrive at the screen $y$ should be the sum of the probabilities to go by way of the upper or lower slit. But in quantum theory, probabilities are squares of amplitudes and

$$|\psi_U(y) + \psi_L(y)|^2 \neq |\psi_U(y)|^2 + |\psi_L(y)|^2.$$ 

Probabilities can only be predicted for sets of alternative coarse-grained histories for which the quantum interference is negligible between every pair of coarse-grained histories in the set (decoherence).

Usual quantum mechanics is not the only way of implementing the three elements of generalized quantum theory. Below we sketch a sum-over-histories generalized quantum theory of space-time. The fine-grained histories are the set of 4D cosmological space-times with matter fields on them. A coarse graining is a partition of this set into (diffeomorphism invariant) classes. A natural measure of interference is described. This is a fully 4D quantum theory without an equivalent 3+1 formulation in terms of states on space-like surfaces and their unitary evolution between them. Rather, the usual 3+1 formulation is emergent for those situations, and for those coarse-grainings, where space-time geometry behaves approximately classically. The intent of this development is not to propose a new quantum theory of gravity. This essentially low energy theory suffers from the usual ultraviolet difficulties. Rather, it is to employ this theory as a model to discuss how quantum mechanics can be generalized to deal with quantum geometry. A common expectation is that space-time is itself emergent from something more fundamental. In that case a generalization of usual quantum mechanics will surely be needed and generalized quantum theory can provide a framework for discovering it.
4.3.4.1 Quantum Mechanics Today

Three features of quantum theory are striking from the present perspective: its success, its rejection by some of the deepest thinkers, and the absence of compelling alternatives [Hartle (2005b)]. Firstly, quantum mechanics must be counted as one of the most successful of all physical theories. Within the framework it provides, a truly vast range of phenomena can be understood and that understanding is confirmed by precision experiment. We perhaps have little evidence for peculiarly quantum phenomena on large and even familiar scales, but there is no evidence that all the phenomena we do see, from the smallest scales to the largest of the universe, cannot be described in quantum mechanical terms and explained by quantum mechanical laws. Indeed, the frontier to which quantum interference is confirmed experimentally is advancing to ever larger, more ‘macroscopic’ systems.\footnote{For an insightful and lucid review see Leggett (2002).}

The textbook electron two-slit experiment has been realized in the laboratory [Tonomura et al. (1989)]. Interference has been confirmed for the biomolecule tetraphenylporphyrin (C_{44}H_{30}N_4) and the fluor fullerene (C_{60}F_{48}) in analogous experiments [Hackermüller et al. (2003)]. Experiments with superconducting squids have demonstrated the coherent superposition of macroscopic currents [Van der Wal et al. (2000); Chiorescu et al. (2003); Friedman et al. (2003)]. In particular, the experiment of Friedman, et al. [Friedman et al. (2003)] exhibited the coherent superposition of two circulating currents whose magnetic moments were of order \(10^{10}\mu_B\) (where \(\mu_B = e\hbar/2m_e c\) is the Bohr magneton). Experiments under development will extend the boundary further [Marshall et al. (2003)]. Experiments of increasing ingenuity and sophistication have extended the regime in which quantum mechanics has been tested. No limit to its validity has yet emerged.

Secondly, even while acknowledging its undoubted empirical success, many of the greatest minds have rejected quantum mechanics as a framework for fundamental theory. Among the pioneers, the names of Einstein, Schrödinger, DeBroglie, and Bohm stand out in this regard. Among our distinguished contemporaries, Adler, Leggett, Penrose, and 't Hooft could probably be counted in this category. Much of this thought has in common the intuition that quantum mechanics is an effective approximation of a more fundamental theory built on a notion of reality closer to that classical physics.

Finally and remarkably, despite eighty years of unease with its basic
premises, and despite having been tested only in a limited, largely microscopic, domain, no fully satisfactory alternative to quantum theory has emerged. By fully satisfactory we mean not only consistent with existing experiment, but also incorporating other seemingly secure parts of modern physics such as special relativity, field theory, and the standard model of elementary particle interactions. As S. Weinberg summarized the situation, “It is striking that it has not so far been possible to find a logically consistent theory that is close to quantum mechanics other than quantum mechanics itself” [Weinberg (1992)]. Alternatives to quantum theory meeting the above criteria would be of great interest if only to guide experiment.

There are several directions under investigation today which aim at a theory from which quantum mechanics would be emergent. For an encyclopedic survey of different interpretations and alternatives to quantum mechanics, see [Auletta (2000)].

Bohmian mechanics [Bohm and Hiley (1993)] in its most representative form is a deterministic but highly non–classical theory of particle dynamics whose statistical predictions largely coincide with quantum theory [Hartle (2004c)]. Fundamental noise [Percival (1998)] or spontaneous dynamical collapse of the wave function [Bassi and Ghirardi (2003); Dowker and Henson (2004)] are the underlying ideas of another class of model theories whose predictions are distinguishable from those of quantum theory, in principle. S. Adler has proposed a statistical mechanics of deterministic matrix models from which quantum mechanics is emergent [Adler (2004)]. G. ‘t Hooft has a different set of ideas for a determinism beneath quantum mechanism that are explained in [‘t Hooft (2006)]. R. Penrose has championed a role for gravity in state vector reduction [Penrose (2000); Penrose (2004)]. This has not yet developed into a detailed alternative theory, but has suggested experimental situations in which the decay of quantum superpositions could be observed [Penrose (2004); Marshall et al. (2003)].

In the face of an increasing domain of confirmed predictions of quantum theory and the absence as yet of compelling alternatives, it seems natural to extend quantum theory as far as it will go, to the largest scales of the universe and the smallest of quantum gravity [Hartle (2005b)].

4.3.4.2 Spacetime and Quantum Theory

Usual, textbook quantum theory incorporates definite assumptions about the nature of space and time. These assumptions are readily evident in the
two laws of evolution for the quantum state $\Psi$. The Schrödinger equation describes its unitary evolution between measurements,

$$i\hbar \frac{\partial}{\partial t} \Psi = H\Psi.$$  \hfill (4.61)

At the time of an ideal measurement, the state is projected on the outcome and renormalized

$$\Psi \rightarrow \frac{P\Psi}{\|P\Psi\|}. \hfill (4.62)$$

The Schrödinger equation (4.61) assumes a fixed notion of time. In the non–relativistic theory, $t$ is the absolute time of Newtonian mechanics. In the flat space–time of special relativity, it is the time of any Lorentz frame. Thus, there are many times but results obtained in different Lorentz frames, are unitarily equivalent. The projection in the second law of evolution (4.62) is in Hilbert space. But in field theory or particle mechanics, the Hilbert space is constructed from configurations of fields or position in physical space. In that sense it is the state on a space–like surface that is projected (4.62).

Because quantum theory incorporates notions of space and time, it has changed as our ideas of space and time have evolved. The accompanying table briefly summarizes this co–evolution. It is possible to view this evolution as a process of increasing generalization of the concepts in the usual theory. Certainly the two laws of evolution (4.61) and (4.62) have to be generalized somehow if space–time geometry is not fixed. One such generalization is offered in this section, but there have been many other ideas [Kuchař (1992)]. And if space–time geometry is emergent from some yet more fundamental description, we can certainly expect that a further generalization, free of any reference to space–time, will be needed to describe that emergence [Hartle (2005b)].

4.3.4.3 The Quantum Mechanics of Closed Systems

Here we briefly review the elements of the modern quantum mechanics of closed systems aimed at a quantum mechanics for cosmology. See, e.g., Griffiths (2002) Omnès (1994) Gell-Mann (1994) for the classic expositions at length or Hartle (1993) for a shorter summary for quantum mechanics of closed systems. To keep the present discussion manageable we focus on a simple model universe of particles moving in a very large box. Everything is contained within the box, in particular galaxies, stars,
Quantum Leap

planets, observers and observed (if any), measured subsystems, and the apparatus that measures them. We assume a fixed background space–time supplying well–defined notions of time. The usual apparatus of Hilbert space, states, operators, Feynman paths, etc. can then be employed in a quantum description of the contents of the box. The essential theoretical inputs to the process of prediction are the Hamiltonian \( H \) and the initial quantum state \( |\Psi\rangle \), the wave function of the universe. These are assumed to be fixed and given.

The most general objective of a quantum theory for the box is the prediction of the probabilities of exhaustive sets of coarse–grained alternative time histories of the particles in the closed system. For instance, we might be interested in the probabilities of an alternative set of histories describing the progress of the Earth around the Sun. Histories of interest here are typically very coarse–grained for at least three reasons: They deal with the position of the Earth’s center–of–mass and not with the positions of all the particles in the universe. The center–of–mass position is not specified to arbitrary accuracy, but to the error we might observe it. The center–of–mass position is not specified at all times, but typically at a series of times.

But, as described in the Introduction, not every set of alternative histories that may be described can be assigned consistent probabilities because of quantum interference. Any quantum theory must therefore not only specify the sets of alternative coarse–grained histories, but also give a rule identifying which sets of histories can be consistently assigned probabilities as well as what those probabilities are. In the quantum mechanics of closed systems, that rule is simple: probabilities can be assigned to just those sets of histories for which the quantum interference between its members is negligible as a consequence of the Hamiltonian \( H \) and the initial state \( |\Psi\rangle \).

We now make this specific for our model universe of particles in a box. Three elements specify this quantum theory; to facilitate later discussion, we give these in a space–time sum–over–histories formulation \cite{Hartle2005b}:

1. **Fine-grained histories**: The most refined description of the particles from the initial time \( t = 0 \) to a suitably large final time \( t = T \) gives their position at all times in between, i.e., their Feynman paths. We denote these simply by \( x(t) \).
2. **Coarse–graining**: The general notion of coarse-graining is a partition of the fine–grained paths into an exhaustive set of mutually exclusive classes \( \{c_\alpha\}, \alpha = 1, 2, \ldots \). For instance, we might partition the fine–
grained histories of the center–of–mass of the Earth by which of an exhaustive and exclusive set of position intervals \( \{ \Delta_\alpha \} \), \( \alpha = 1, 2, \ldots \) the center–of–mass passes through at a series of times \( t_1, \ldots, t_n \). Each coarse–grained history consists of the bundle of fine–grained paths that pass through a specified sequence of intervals at the series of times. Each coarse–grained history specifies an orbit where the center–of–mass position is localized to a certain accuracy at a sequence of times.

(3) **Measure of Interference**: Branch state vectors \( | \Psi_\alpha \rangle \) can be defined for each coarse–grained history in a partition of the fine–grained histories into classes \( \{ c_\alpha \} \) as follows

\[
\langle x | \Psi_\alpha \rangle = \int_{c_\alpha} \mathcal{D}[x] \exp(iS[x(t)]/\hbar) \langle x' | \Psi \rangle.
\]

Here, \( S[x(t)] \) is the action for the Hamiltonian \( H \). The integral is over all paths starting at \( x' \) at \( t = 0 \), ending at \( x \) at \( t = T \), and contained in the class \( c_\alpha \). This includes an integral over \( x' \). (For those preferring the Heisenberg picture, this is equivalently

\[
| \Psi_\alpha \rangle = e^{-iHT/\hbar} P^{\alpha_n}_{c_\alpha}(t_n) \cdots P^{\alpha_1}_{c_\alpha}(t_1) | \Psi \rangle
\]

when the class consists of restrictions to position intervals at a series of times and the \( P \)'s are the projection operators representing them.)

The measure of quantum interference between two coarse–grained histories is the overlap of their branch state vectors

\[
D(\alpha', \alpha) \equiv \langle \Psi_{\alpha'} | \Psi_\alpha \rangle.
\]

This is called the **decoherence functional**.

When the interference between each pair of histories in a coarse–grained set is negligible

\[
\langle \Psi_\alpha | \Psi_\beta \rangle \approx 0 \quad \text{all} \quad \alpha \neq \beta,
\]

the set of histories is said to **decohere**. The probability of an individual history in a decoherent set is

\[
p(\alpha) = \| | \Psi_\alpha \rangle \|^2.
\]

The decoherence condition (4.65) is a sufficient condition for the probabilities (4.66) to be consistent with the rules of probability theory. Specifically,

---

\[ ^{10} \] This is the *medium* decoherence condition. For a discussion of other conditions, see, e.g., [Gell-Mann and Hartle (1990); Gell-Mann and Hartle (1995); Hartle (2004a)].
the $p$'s obey the sum rules

$$p(\bar{\alpha}) \approx \sum_{\alpha \in \bar{\alpha}} p(\alpha)$$

(4.68)

where $\{\bar{c}_\alpha\}$ is any coarse-graining of the set $\{c_\alpha\}$, i.e., a further partition into coarser classes. It was the failure of such a sum rule that prevented consistent probabilities from being assigned to the two histories previously discussed in the two-slit experiment (Figure 1). That set of histories does not decohere.

Decoherence of familiar quasi-classical variables is widespread in the universe. Imagine, for instance, a dust grain in a superposition of two positions, a multimeter apart, deep in intergalactic space. The $10^{11}$ cosmic background photons that scatter off the dust grain every second dissipate the phase coherence between the branches corresponding to the two locations on the time scale of about a nanosecond [Joos and Zeh (1985)].

Measurements and observers play no fundamental role in this generalization of usual quantum theory. The probabilities of measured outcomes can, of course, be computed and are given to an excellent approximation by the usual story. But, in a set of histories where they decohere, probabilities can be assigned to the position of the Moon when it is not being observed and to the values of density fluctuations in the early universe when there were neither measurements taking place nor observers to carry them out [Hartle (2005b)].

### 4.3.4.4 Quantum Theory in 3+1 Form

The quantum theory of the model universe in a box in the previous section is in fully 4D space–time form. The fine-grained histories are paths in space–time, the coarse–grainings were partitions of these, and the measure of interference was constructed by space–time path integrals. No mention was made of states on space–like surfaces or their unitary evolution.

However, as originally shown by Feynman [Feynman (1948); Feynman and Hibbs (1965)], this space–time formulation is equivalent to the familiar 3+1 formulation in terms of states on space–like surfaces and their unitary evolution through a foliating family of such surfaces. This section briefly sketches that equivalence emphasizing properties of space–time and the fine–grained histories that are necessary for it to hold.

The key observation is illustrated in Figure 4.3.4.4. Sums–over–histories that are single–valued in time can be factored across constant time surfaces.
Fig. 4.3 The origin of states on a space–like surface. These space–time diagrams are a schematic representation of (4.69). The amplitude for a particle to pass from point $A$ at time $t = 0$ to a point $B$ at $t = T$ is a sum over all paths connecting them weighted by $\exp(iS[x(t)])$. That sum can be factored across an intermediate constant time surface as shown at right into product of a sum from $A$ to $x$ on the surface and a sum from $x$ to $B$ followed by a sum over all $x$. The sums in the product define states on the surface of constant time at $t$. The integral over $x$ defines the inner product between such states, and the path integral construction guarantees their unitary evolution in $t$. Such factorization is possible only if the paths are single valued functions of time (modified and adapted from [Hartle (2005b)]).

A formula expressing this idea is [Hartle (2005b)]:

$$\int_{[A,B]} D[x] e^{iS[x(t)]}/\hbar = \int \psi_B^\ast(x,t) \psi_A(x,t) \, dx.$$  \(4.69\)

The sum on the left is over all paths from $A$ at $t = 0$ to $B$ at $t = T$. The amplitude $\psi_A(x,t)$ is the sum of $\exp\{iS[x(t)]\}$ over all paths from $A$ at $t = 0$ to $x$ at a time $t$ between 0 and $T$. The amplitude $\psi_B(x,t)$ is similarly constructed from the paths between $x$ at $t$ to $B$ at $T$.

The wave function $\psi_A(x,t)$ defines a state on constant time surfaces. Unitary evolution by the Schrödinger equation follows from its path integral construction.\(^\text{11}\) The inner product between states defining a Hilbert space is specified by (4.69). In this way, the familiar 3+1 formulation of quantum mechanics is recovered from its space–time form.

The equivalence represented in (4.69) relies on several special assumptions about the nature of space–time and the fine–grained histories. In particular, it requires:\(^\text{12}\)

\(^{11}\)Reduction of the state vector (4.62) also follows from the path integral construction [Caves (1986)] when histories are coarse–grained by intervals of position at various times.

\(^{12}\)The usual 3+1 formulation is also restricted to coarse–grained histories specified by
Quantum Leap

- A fixed Lorentzian space–time geometry to define time–like and space–like directions.
- A foliating family of space–like surfaces through which states can evolve.
- Fine–grained histories that are single-valued in the time labeling the space–like surfaces in the foliating family.

As an illustrative example where the equivalence does not hold, consider quantum field theory in a fixed background space–time with closed time–like curves (CTCs) such as those that can occur in wormhole space–times [Morris et al. (1988)]. The fine–grained histories are 4D field configurations that are single-valued on space–time. But there is no foliating family of space–like surfaces with which to define the Hamiltonian evolution of a quantum state. Thus, there is no usual 3+1 formulation of the quantum mechanics of fields in space–times with CTCs.

However, there is a 4D sum–over–histories formulation of field theory in space–times with CTCs [Hartle (1994a); Friedman et al. (1992); Rosenberg (1998)]. The resulting theory has some unattractive properties such as acausality and non–unitarity. But it does illustrate how closely usual quantum theory incorporates particular assumptions about space–time, and also how these requirements can be relaxed in a suitable generalization of the usual theory.

4.3.4.5 Generalized Quantum Theory

In generalizing usual quantum mechanics to deal with quantum space–time, some of its features will have to be left behind and others retained. What are the minimal essential features that characterize a quantum mechanical theory? The generalized quantum theory framework [Hartle (1991a); Hartle (1993); Isham (1994)] provides one answer to this question. Just three elements abstracted from the quantum mechanics of closed systems above define a generalized quantum theory [Hartle (2005b)]:

- Fine–grained histories: the sets of alternative fine–grained histories of the closed system which are the most refined descriptions of it physically possible.
Coarse–grained histories: these are partitions of a set of fine–grained histories into an exhaustive set of exclusive classes \( \{ c_\alpha \} \), \( \alpha = 1, 2, \ldots \). Each class is a coarse–grained history.

Decoherence functional: the measure of quantum interference \( D(\alpha, \alpha') \) between pairs of histories in a coarse–grained set, meeting the following conditions:

(i) Hermiticity: \( D(\alpha, \alpha') = D^*(\alpha', \alpha) \)

(ii) Positivity: \( D(\alpha, \alpha) \geq 0 \)

(iii) Normalization: \( \Sigma_{\alpha, \alpha'} D(\alpha, \alpha') = 1 \)

(iv) Principle of superposition: If \( \{ \bar{c}_\bar{\alpha} \} \) is a further coarse–graining of \( \{ c_\alpha \} \), then

\[
D(\bar{\alpha}, \bar{\alpha}') = \sum_{\alpha \in \bar{\alpha}} \sum_{\alpha' \in \bar{\alpha}'} D(\alpha, \alpha')
\]

Probabilities \( p(\alpha) \) are assigned to sets of coarse–grained histories when they decohere according to the basic relation

\[
D(\alpha, \alpha') \approx \delta_{\alpha \alpha'} p(\alpha).
\] (4.70)

These \( p(\alpha) \) satisfy the basic requirements for probabilities as a consequence of (i)–(iv) above. In particular, they satisfy the sum rule

\[
p(\bar{\alpha}) = \sum_{\alpha \in \bar{\alpha}} p(\alpha)
\] (4.71)

as a consequence of (i)–(iv) and decoherence. For instance, the probabilities of an exhaustive set of alternatives always sum to 1.

The sum–over–histories formulation of usual quantum mechanics given above is a particular example of a generalized quantum theory. The decoherence functional \( (4.63) \) satisfies the requirements (i)–(iv). But its particular form is not the only way of constructing a decoherence functional. Therein lies the possibility of generalization.

4.3.4.6 A Quantum Theory of Spacetime Geometry

The low energy, effective theory of quantum gravity is a quantum version of general relativity with a space–time metric \( g_{\alpha\beta}(x) \) coupled to matter fields. Of course, the divergences of this effective theory have to be regulated to extract predictions from it (perhaps, most naturally by discrete approximations to geometry such as the Regge calculus, see, e.g. [Hartle (1985a)])
These predictions can therefore be expected to be accurate only for limited coarse-grainings and certain states. But this effective theory does supply an instructive model for generalizations of quantum theory that can accommodate quantum space–time. This generalization is sketched in this section.

The key idea is that the fine-grained histories do not have to represent evolution in space–time. Rather they can be histories of space–time. For this discussion we take these histories to be spatially closed cosmological 4-geometries represented by metrics $g_{\alpha\beta}(x)$ on a fixed manifold $M = \mathbb{R} \times M^3$ where $M^3$ is a closed 3-manifold. For simplicity, we restrict attention to a single scalar matter field $\phi(x)$.

The three ingredients of a generalized quantum theory for space–time geometry are then as follows [Hartle (2005b)]:

- **Fine-grained Histories**: A fine-grained history is defined by a 4D metric and matter field configuration on $M$.
- **Coarse-grainings**: The allowed coarse-grainings are partitions of the metrics and matter fields into 4D diffeomorphism invariant classes $\{c_\alpha\}$.
- **Decoherence Functional**: A decoherence functional constructed on sum-over-history principles analogous to that described above for usual quantum theory. Schematically, branch state vectors $|\Psi_\alpha\rangle$ can be constructed for each coarse-grained history by summing over the metrics and fields in the corresponding class $c_\alpha$ of fine-grained histories [Hartle (2005b)]

$$|\Psi_\alpha\rangle = \int_{c_\alpha} D[g] D[\phi] \exp\left\{ iS[g, \phi]/\hbar \right\} |\Psi\rangle.$$ (4.72)

A decoherence functional satisfying the requirements above is

$$D(\alpha', \alpha) = \langle \Psi_{\alpha'} | \Psi_\alpha \rangle.$$ (4.73)

Here, $S[g, \phi]$ is the action for general relativity coupled to the field $\phi(x)$, and $|\Psi\rangle$ is the initial cosmological state. The construction is only schematic because we did not spell out how the functional integrals are defined or regulated, nor did we specify the product between states that is implicit in both (4.72) and (4.73). These details can be made specific in models [Hartle (1995b), Hartle and Marolf (1997), Craig and Hartle (2004)], but they will not be needed for the subsequent discussion.

A few remarks about the coarse-grained histories may be helpful. To every physical assertion that can be made about the geometry of the uni-
verse and the fields within, there corresponds a diffeomorphism invariant partition of the fine-grained histories into the class where the assertion is true and the class where it is false. The notion of coarse-grained history described above therefore supplies the most general notion of alternative describable in space-time form. Among these we do not expect to find local alternatives because there is no diffeomorphism invariant notion of locality. In particular, we do not expect to find alternatives specified at a moment of time. We do expect to find alternatives referring to the kind of relational observables discussed in [Giddings et al. (2006)] and the references therein. We also expect to find observables referring to global properties of the universe such as the maximum size achieved over the history of its expansion.

This generalized quantum mechanics of space-time geometry is in fully space-time form with alternatives described by partitions of 4D histories and a decoherence functional defined by sums over those histories. It is analogous to the space-time formulation of usual quantum theory reviewed above.

However, unlike the above theory, we cannot expect an equivalent 3+1 formulation, of the kind described above, expressed in terms of states on space-like surfaces and their unitary evolution between these surfaces. The fine-grained histories are not ‘single-valued’ in any geometrically defined variable labelling a space-like surface. They therefore cannot be factored across a space-like surface as in (4.69). More precisely, there is no geometrical variable that picks out a unique space-like surface in all geometries.

Even without a unitary evolution of states the generalized quantum theory is fully predictive because it assigns probabilities to the most general sets of coarse-grained alternative histories described in space-time terms when these are decoherent.

How then is usual quantum theory used every day, with its unitarily evolving states, connected to this generalized quantum theory that is free from them? The answer is that usual quantum theory is an approximation to the more general framework that is appropriate for those coarse-grainings and initial state $|\Psi\rangle$ for which space-time behaves classically. One equation will show the origin of this relation. Suppose we have a coarse-

\[13\text{Spacelike surfaces labelled by the trace of the extrinsic curvature } K \text{ foliate certain classes of classical space-times obeying the Einstein equation [Marsden and Tipler (1980)]. However, there is no reason to require that non-classical histories be foliable in this way. It is easy to construct geometries where surfaces of a given } K \text{ occur arbitrarily often.}\]
graining that distinguishes between fine–grained geometries only by their behavior on scales well above the Planck scale. Then, for suitable states $|\Psi\rangle$ we expect that the integral over metrics in (4.73) can be well approximated semiclassically by the method of steepest descents. Suppose further for simplicity that only a single classical geometry with metric $\hat{g}_{\alpha\beta}$ dominates the semiclassical approximation. Then, (4.73) becomes

$$|\Psi_{\alpha}\rangle \approx \int \mathcal{D}[\phi] \exp\left\{iS[\hat{g},\phi]/\hbar\right\} |\Psi\rangle,$$  \hspace{1cm} (4.74)

where $\hat{c}_{\alpha}$ is the coarse–graining of $\phi(x)$ arising from $c_{\alpha}$ and the restriction of $g_{\alpha\beta}(x)$ to $\hat{g}_{\alpha\beta}(x)$. Equation (4.74) effectively defines a quantum theory of the field $\phi(x)$ in the fixed background space–time with the geometry specified by $\hat{g}_{\alpha\beta}(x)$. This is familiar territory. Field histories are single valued on space–time. This sum–over–fields can thus be factored across space–like surfaces in the geometry $\hat{g}$ as in (4.69) to define field states on space–like surfaces, their unitary evolution, and their Hilbert space product. Usual quantum theory is thus recovered when space–time behaves classically and provides the fixed space–time geometry on which usual quantum theory relies. From this perspective, familiar quantum theory and its unitary evolution of states is an effective approximation to a more general sum–over–histories formulation of quantum theory. The approximation is appropriate for those coarse-grainings and initial states in which space–time geometry behaves classically.

4.3.4.7 Beyond Spacetime

The generalized quantum theory of space–time sketched in the previous section assumed that geometry was a fundamental variable — part of the description of the fine–grained histories. But on almost every frontier in quantum gravity one finds the idea that continuum geometry is not fundamental, but will be replaced by something more fundamental. This is true for string theory [Seiberg (2006)], loop quantum gravity [Ashtekar and Lewandowski (2004)], and the causal set program [Dowker (2005); Henson (2006)] although space does not permit a review of these speculations.

Can generalized quantum theory serve as a framework for theories where space–time is emergent rather than fundamental? Certainly we cannot expect to have a notion of ‘history’. But we can expect some fine–grained description, or a family of equivalent ones, and that is enough. A general-
ized quantum theory needs [Hartle (2005b)]:

- The possible fine-grained descriptions of the system.
- The coarse-grained descriptions constructed from the fine-grained ones.
- A measure of quantum interference between different coarse-grained descriptions respecting conditions (i)–(iv) above.

Generalized quantum theory requires neither space nor time and can therefore serve as the basis for a quantum theory in which space-time is emergent.

4.3.4.8 Emergence/Excess Baggage

The word ‘emergent’ appears in a number of places in the previous discussion. It probably has many meanings. This section aims at a more precise understanding of what is meant by the term in this essay.

Suppose we have a quantum theory defined by certain sets of fine-grained histories, coarse-grainings, and a decoherence functional. Let’s call this the fundamental theory. It may happen that the decoherence and probabilities of limited kinds of sets of coarse-grained histories are given approximately by a second, effective theory. The two theories are related in the following way [Hartle (2005b)]:

- Every fine-grained history of the effective theory is a coarse-grained history of the fundamental theory.
- The decoherence functionals approximately agree on a limited class of sets of coarse-grained histories.

\[ D_{\text{fund}}(\alpha', \alpha) \approx D_{\text{eff}}(\alpha', \alpha). \] (4.75)

On the right, \( \alpha' \) and \( \alpha \) refer to the fine-grained histories of the effective theory. On the left, they refer to the corresponding coarse-grained histories of the fundamental theory.

When two theories are related in this way we can say that the effective theory is emergent from the fundamental theory. Loosely we can say that the restrictions, and the concepts that characterize them, are emergent. It should be emphasized that an approximate equality like [4.75] can be expected to hold, not just as a consequence of the particular dynamics incorporated into decoherence functionals, but also only for particular states.
Several examples of emergence in this sense have been considered in this essay: There is the possible emergence of a generalized quantum theory of space-time geometry from a theory in which space-time is not fundamental. There is the emergence of a 3+1 quantum theory of fields in a fixed background geometry from a 4D generalized quantum theory in which geometry is a quantum variable. There is the emergence of the approximate quantum mechanics of measured subsystems (textbook quantum theory) from the quantum mechanics of the universe. And there is the emergence of classical physics from quantum physics.

Instead of looking at an effective theory as a restriction of a more fundamental one, we may look at the fundamental theory as a generalization of the effective one. That perspective is important because generalization is a way of searching for more comprehensive theories of nature. In passing from the specific to the more general some ideas have to be discarded. They are often ideas that were once perceived to be general because of our special place in the universe and the limited range of our experience. But, in fact, they arise from special situations in a more general theory. They are ‘excess baggage’ that has to be discarded to reach a more comprehensive theory [Hartle (1990)]. Emergence and excess baggage are two ways of looking at the same thing [Hartle (2005b)].

4.3.4.9 Emergence of Signature

Classical space–time has Lorentz signature. At each point it is possible to choose one time–like direction and three orthogonal space–like ones. There are no physical space–times with zero time–like directions or with two time–like directions. But is such a seemingly basic property fundamental, or is it rather, emergent from a quantum theory of space–time which allows for all possible signatures? This section sketches a simple model where that happens.

Classical behavior requires particular states [Hartle (1994b)]. Let’s consider the possible classical behaviors of cosmological geometry assuming the no–boundary quantum state of the universe [Hartle and Hawking (1983)] in a theory with only gravity and a cosmological constant $\Lambda$. The no–boundary wave function is given by a sum–over–geometries of the schematic form [Hartle (2005b)]

$$\Psi [h] = \int_c D[g] e^{-I[g]/\hbar}.$$  \hspace{1cm} (4.76)
For simplicity, we consider a fixed manifold $M$. The key requirement is that it be compact with one boundary for the argument of the wave function and no other boundary. The functional $I[g]$ is the Euclidean action for metric defining the geometry on $M$. The sum is over a complex contour $C$ of $g$’s that have finite action and match the three-metric $h$ on the boundary that is the argument of $\Psi$.

Quantum theory predicts classical behavior when it predicts high probability for histories exhibiting the correlations in time implied by classical deterministic laws [Gell-Mann and Hartle (1993); Hartle (1994b)]. The state $\Psi$ is an input to the process of predicting those probabilities as described above. However, plausibly the output for the predicted classical space–times in this model are the extrema of the action in (4.76). We will assume this (see [Hartle (1995b)] for some justification). Further, to keep the discussion manageable, we will restrict it to the real extrema. These are the real tunneling geometries discussed in a much wider context in [Gibbons and Hartle (1990)].

Fig. 4.4 The emergence of the Lorentz signature $(-, +, +, +)$ of space–time. The semi-classical geometry describing a classical space–time which becomes large according to the ‘no-boundary’ proposal for the universe’s quantum state. The model is pure gravity and a cosmological constant. Purely Euclidean geometries $(+, +, +, +)$ or purely Lorentzian geometries are not allowed as described in the text. What is allowed is the real tunneling geometry illustrated above consisting of half a Euclidean four-sphere joined smoothly onto an expanding Lorentzian de Sitter space at the moment of maximum contraction. This can be described as the nucleation of classical Lorentzian space–time. There is no similar nucleation of a classical geometry with signature $(-, -, +, +)$ because it could not match the Euclidean one across a space–like surface (modified and adapted from [Hartle (2005b)]).

$^{14}$Even the notion of manifold may be emergent in a more general theory of certain complexes [Hartle (1985b); Schleich and Witt (1993)].
Let us ask for the semiclassical geometries which become large, i.e., contain symmetric three surfaces with size much larger than \((1/\Lambda)^{1/2}\). There are none with Euclidean signature. The purely Euclidean extremum is the round 4–sphere with linear size \((1/\Lambda)^{1/2}\) and contains no symmetric three surfaces with larger size. There are none with purely Lorentzian signature either because these cannot be regular on \(M\). There are, however, tunneling solutions of the kind illustrated in Figure 4.3.4.9 in which half of a Euclidean 4–sphere is matched to expanding DeSitter space across a surface of vanishing extrinsic curvature. Could a space–time with two time and two space directions be nucleated in this way? The answer is ‘no’ because the geometry on a surface could not have the three space–like directions necessary to match onto the half of a 4–sphere. Thus, in this very simple model, with many assumptions, if we live in a large universe it must have one time and three space dimensions. The Lorentzian signature of classical space–time is an emergent property from an underlying theory not committed to this signature \cite{Hartle2005b}.

4.3.4.10 Beyond Quantum Theory

The path of generalization in the previous sections began with the textbook quantum mechanics of measurement outcomes in a fixed space–time and ended in a quantum theory where neither measurements nor space–time are fundamental. In this journey, the principles of generalized quantum theory are preserved, in particular the idea of quantum interference and the linearity inherent in the principle of superposition. But the end of this path is strikingly different from its beginning.

The founders of quantum theory thought that the indeterminacy of quantum theory “reflected the unavoidable interference in measurement dictated by the magnitude of the quantum of the action” (Bohr). But what then is the origin of quantum indeterminacy in a closed quantum universe which is never measured? Why enforce the principle of superposition in a framework for prediction of the universe which has but a single quantum state? In short, the endpoint of this journey of generalization forces us to ask J. Wheeler’s famous question, “How come the quantum?” \cite{Wheeler1986}.

Could quantum theory itself be an emergent effective theory? Many have thought so. Extending quantum mechanics until it breaks could be one route to finding out. If space–time geometry is not fundamental, quantum
mechanics will need further generalization and generalized quantum theory provides one framework for exploring that. For more details, see [Hartle (2005b)].

4.3.5 Anthropic String Landscape

![Diagram of a scalar field φ in a false vacuum.](image)

Fig. 4.5 A scalar field $\phi$ in a false vacuum. Note that the potential energy $V(\phi)$ is higher than that in the true vacuum or ground state, but there is a barrier preventing the field from classically rolling down to the true vacuum. Therefore, the transition to the true vacuum must be stimulated by the creation of high energy particles or through quantum mechanical tunnelling.

The **anthropic string landscape** refers to the large number of different false vacua in string theory. Recall that a false vacuum (see Figure 4.5) is a metastable sector of a quantum field theory (QFT) which appears to be a perturbative vacuum but is unstable to instanton effects which tunnel to a lower energy state. This tunnelling\(^{15}\) can be caused by quantum fluctuation.

\(^{15}\)Recall that tunneling is the quantum–mechanical effect of transitioning through a classically–forbidden energy state. It can be generalized to other types of classically–forbidden transitions as well. For example, consider rolling a ball up a hill. If the ball is not given enough velocity, then it will not roll over the hill. This scenario makes sense from the standpoint of classical mechanics, but is an inapplicable restriction in quantum mechanics simply because quantum mechanical objects do not behave like classical objects such as balls. On a quantum scale, objects exhibit wavelike behavior. For a quantum particle moving against a potential energy 'hill', the wave $\psi$–function describing the particle can extend to the other side of the hill. This wave represents the probability of finding the particle in a certain location, meaning that the particle has the possibility of being detected on the other side of the hill. This behavior is called tunnelling; it is as if the particle has 'dug' through the potential hill. As this is a quantum and non–classical effect, it can generally only be seen in nanoscopic phenomena, where the wave behavior of particles is more pronounced. Availability of states is necessary for tunnelling to occur. In the above example, the quantum mechanical ball will not appear inside the hill because there is no available 'space' for it to exist, but it can tunnel to the other side of the hill, where there is free space. Analogously, a particle can tunnel...
tions or the creation of high energy particles. Simply put, the false vacuum is a state of a physical theory which is not the lowest energy state, but is nonetheless stable for some time. This is analogous to metastability for the first order phase transitions.

In a physical theory in a false vacuum, the system moves to a lower through the barrier, but unless there are states available within the barrier, the particle can only tunnel to the other side of the barrier. The wave–function describing a particle only expresses the probability of finding the particle at a location assuming a free state exists.

Recall that in thermodynamics, phase transition is the transformation of a thermodynamic system from one phase to another. The distinguishing characteristic of a phase transition is an abrupt sudden change in one or more physical properties, in particular the heat capacity, with a small change in a thermodynamic variable such as the temperature. The first–order phase transitions are those that involve a latent heat. During such a transition, a system either absorbs or releases a fixed (and typically large) amount of energy. Because energy cannot be instantaneously transferred between the system and its environment, first–order transitions are associated with ‘mixed–phase regimes’, in which some parts of the system have completed the transition and others have not. This phenomenon is familiar to anyone who has boiled a pot of water: the water does not instantly turn into gas, but forms a turbulent mixture of water and water vapor bubbles. Mixed–phase systems are difficult to study, because their dynamics are violent and hard to control. However, many important phase transitions fall in this category, including the solid/liquid/gas transitions and Bose–Einstein condensation.

On the other hand, the second–order phase transitions are the continuous phase transitions, such as the ferromagnetic transition and the superfluid transition. They have no associated latent heat.

In systems containing liquid and gaseous phases, there exist a special combination of pressure and temperature, known as the critical point, at which the transition between liquid and gas becomes a second–order transition. Near the critical point, the fluid is sufficiently hot and compressed that the distinction between the liquid and gaseous phases is almost non–existent. Phase transitions often (but not always) take place between phases with different symmetry. Consider, for example, the transition between a fluid (i.e., liquid or gas) and a crystalline solid. A fluid, which is composed of atoms arranged in a disordered but homogeneous manner, possesses continuous translational symmetry: each point inside the fluid has the same properties as any other point. A crystalline solid, on the other hand, is made up of atoms arranged in a regular lattice. Each point in the solid is not similar to other points, unless those points are displaced by an amount equal to some lattice spacing.

Generally, we may speak of one phase in a phase transition as being more symmetrical than the other. The transition from the more symmetrical phase to the less symmetrical one is a symmetry–breaking process. In the fluid–solid transition, for example, we say that continuous translation symmetry is broken.

The ferromagnetic transition is another example of a symmetry–breaking transition, in this case the symmetry under reversal of the direction of electric currents and magnetic field lines. This symmetry is referred to as up–down symmetry or time–reversal symmetry. It is broken in the ferromagnetic phase due to the formation of magnetic domains containing aligned magnetic moments. Inside each domain, there is a magnetic field pointing in a fixed direction chosen spontaneously during the phase transition. The name ‘time–reversal symmetry’ comes from the fact that electric currents reverse direction when the time coordinate is reversed.
energy state (either the true vacuum, or another, lower energy vacuum) through a process known as bubble nucleation \cite{Coleman (1977)}. In this, instanton effects cause a bubble to appear in which fields have their true vacuum values inside. Therefore, the interior of the bubble has a lower energy. The walls of the bubble (aka domain walls) have a surface tension, as energy is expended as the fields roll over the potential barrier to the lower energy vacuum. The most likely size of the bubble is determined in the semiclassical approximation\textsuperscript{17} to be such that the bubble has zero 

\textsuperscript{17}Recall that the semiclassical approximation may refer to quantum-mechanical calculations that are obtained by considering a small perturbation of a classical calculation, for example the WKB approximation in non–relativistic quantum mechanics or the loop expansion or the instanton methods in quantum field theory. In quantum field theory, a semiclassical correction arises from one–loop Feynman diagrams. The semiclassical effective action is \[ \Gamma[\phi] = S[\phi] + \frac{1}{2} \text{Tr} \left[ \ln S^{(2)}[\phi] \right] + ... \]
total change in the energy: the decrease in energy by the true vacuum in the interior is compensated by the tension of the walls. Any increase in size of the bubble will decrease its potential energy, as the energy of the wall increases as the area of a sphere $4\pi r^2$ but the negative contribution of the interior increases more quickly, as the volume of a sphere $4/3\pi r^3$. Therefore, after the bubble is nucleated, it quickly begins expanding at very nearly the speed of light. The excess energy contributes to the very large kinetic energy of the walls. If two bubbles are nucleated and they eventually collide, it is thought that particle production occurs where the walls collide.

The tunnelling rate is increased by increasing the energy difference between the two vacua and decreased by increasing the height or width of the barrier.

The addition of gravity to the story leads to a considerably richer variety of phenomena. The key insight is that a false vacuum with positive potential energy density is a de Sitter vacuum, in which the potential energy acts as a cosmological constant and the universe is undergoing the exponential expansion of de Sitter space. This leads to a number of interesting effects,

18 Recall that nD de Sitter space (see, e.g., Coxeter (1943)) is the maximally symmetric, simply-connected, Lorentzian manifold with constant positive curvature. It may be regarded as the Lorentzian analog of an $n$–sphere (with its canonical Riemannian metric). De Sitter space is most easily defined as a submanifold of Minkowski space in one higher dimension. Take Minkowski space $\mathbb{R}^{1,n}$ with the standard metric,

$$ds^2 = -dx_0^2 + \sum_{i=1}^{n} dx_i^2,$$

De Sitter space is the submanifold described by the hyperboloid,

$$-x_0^2 + \sum_{i=1}^{n} x_i^2 = \alpha^2,$$

where $\alpha$ is some positive constant with dimensions of length. The metric on de Sitter space is the metric induced from the ambient Minkowski metric. One can check that the induced metric is nondegenerate and has Lorentzian signature. Topologically, a simply–connected de Sitter space is $\mathbb{R} \times S^{n-1}$. The isometry group of de Sitter space is the Lorentz group $O(1,n)$. The metric therefore then has $n(n+1)/2$ independent Killing vectors and is maximally symmetric. Every maximally symmetric space has constant curvature. The Riemann curvature tensor of de Sitter space is given by

$$R_{\rho\sigma\mu\nu} = \frac{1}{\alpha^2}(g_{\rho\mu}g_{\sigma\nu} - g_{\rho\nu}g_{\sigma\mu}).$$

De Sitter space is an Einstein manifold since the Ricci tensor is proportional to the metric

$$R_{\mu\nu} = \frac{n-1}{\alpha^2} g_{\mu\nu}.$$
first studied by Coleman and de Luccia [Coleman and Luccia (1980)].

(1) Tunnelling from a space with zero potential energy (e.g., Minkowski space) to negative potential energy: the walls of the bubble grow at the speed of light, as described above. However, the interior of the bubble rapidly collapses, as anti–de Sitter space and the universe ends (see ultimate fate of the universe and vacuum metastability event, below).

(2) Tunnelling from a space of positive potential energy (de Sitter space) to one of vanishing potential energy (Minkowski space). In this case, the volume of the bubble continues to grow at the speed of light. Since the exterior of the bubble is expanding exponentially, however, and the Minkowski space is not, unlike the non–gravitational case, the whole of space time need never be dominated by the lower energy vacuum. If the tunnelling rate is slow enough, the exponentially expanding space in the false vacuum state can expand quickly enough that the bubbles of lower–energy space never begin to collide and convert all of space time to the lower energy state. That is, the tunnelling is competing with rapid expansion, and the exponential expansion can be rapid enough that the tunnelling effect is overwhelmed.

In the language of general relativity, de Sitter space is the maximally symmetric, vacuum solution of Einstein’s field equation with a positive cosmological constant given by:

$$\Lambda = \frac{(n-1)(n-2)}{2\alpha^2}.$$  

The scalar curvature of de Sitter space is given by

$$R = \frac{n(n-1)}{\alpha^2} = \frac{2n}{n-2} \Lambda.$$  

We can introduce static coordinates for de Sitter space as

$$x_0 = \sqrt{\alpha^2 - r^2 \sinh(t/\alpha)}, \quad x_1 = \sqrt{\alpha^2 - r^2 \cosh(t/\alpha)}, \quad x_i = rz_i \quad (2 \leq i \leq n),$$

where $z_i$ gives the standard embedding of the $(n-2)$–sphere in $\mathbb{R}^{n-1}$. In these coordinates the de Sitter metric takes the form

$$ds^2 = -\left(1 - \frac{r^2}{\alpha^2}\right)dt^2 + \left(1 - \frac{r^2}{\alpha^2}\right)^{-1}dr^2 + r^2 d\Omega_{n-2}^2.$$  

When $n = 4$, de Sitter space is considered to be a cosmological model for the physical universe, called de Sitter universe. In this case, we have $\Lambda = 3/\alpha^2$ and $R = 4\Lambda = 12/\alpha^2$.

Many inflationary models are approximately de Sitter space and can be modelled by giving the Hubble parameter a mild time dependence. For simplicity, some calculations involving inflation in the early universe can be performed in de Sitter space rather than a more realistic inflationary universe. By using the de Sitter universe instead, where the expansion is truly exponential, there are many simplifications.
(3) Tunnelling from positive potential energy to lower, positive potential energy. Just as for the above case, the more rapid exponential expansion of the higher energy false vacuum can continue to dominate.

(4) Tunnelling from positive potential energy to negative potential energy. This effect is highly suppressed: the expansion of the positive energy vacuum dominates the contraction of the negative energy vacuum.

A final kind of tunnelling is the Hawking–Moss instanton [Hawking and Moss (1982)]. This occurs when the size of the Coleman–de Luccia bubble is larger than the size of the universe, in a closed universe, or of the horizon. In this case, the entire universe tunnels from the false vacuum to the true vacuum at once.

In his original proposal for cosmic inflation [Guth (1981)], A. Guth proposed that inflation could end through quantum mechanical bubble nucleation of the sort described above. It was soon understood that a homogeneous and isotropic universe could not be preserved through the violent tunnelling process. This led A. Linde [Linde (1982)] and, independently, A. Albrecht and P. Steinhardt [Albrecht and Steinhardt (1982)] to propose the so-called ‘new inflation’ or slow-roll inflation, in which no tunnelling occurs, and the inflationary scalar field instead rolls down a gentle slope. A more recent application of these tunnelling phenomena in cosmology and particle physics is the string landscape in which string theory is conjectured to be populated by an exponentially large ‘discretuum’ of false vacua, and the small observed value of the cosmological constant (see dark energy) can be explained by the anthropic principle and quantum mechanical tunnelling to the lowest positive energy vacuum.

In their paper, Coleman and de Luccia noted [Coleman and Luccia (1980)]: “The possibility that we are living in a false vacuum has never been a cheering one to contemplate. Vacuum decay is the ultimate ecological catastrophe; in the new vacuum there are new constants of nature; after vacuum decay, not only is life as we know it impossible, so is chemistry as we know it. However, one could always draw stoic comfort from the possibility that perhaps in the course of time the new vacuum would sustain, if not life as we know it, at least some structures capable of knowing joy. This possibility has now been eliminated.” The possibility that we are living in a false vacuum has been considered. If a bubble of lower energy vacuum were nucleated, it would approach at nearly the speed of light and destroy the Earth instantaneously, without any forewarning. Thus, this vacuum metastability event is a theoretical doomsday event.
The string landscape arises from the idea that there are an extremely large number of metastable vacua (ground states) in string theory (see, e.g., [Douglas (2003)])]. The large number of possibilities arise from different choices of Calabi–Yau manifolds and different values of generalized magnetic fluxes over different homology cycles. This large number of de–Sitter like metastable vacua [Bousso and Polchinski (2000)] is thought by some physicists to be large enough that the known laws of physics, the Standard Model and general relativity with a positive cosmological constant, occurs in at least one, although computing quantities such as masses of particles and Yukawa couplings for even a single vacuum is a technically difficult problem. The problem of enumerating all the vacua is thought to be NP complete [Denef and Douglas (2006)].

4.3.5.1 The Landscape

In this subsection, following [Susskind (2003)], we now give a detailed description of the anthropic string landscape. Recall that the world–view shared by most physicists is that the laws of nature are uniquely described by some special action principle that completely determines the vacuum, the spectrum of elementary particles, the forces and the symmetries. Experience with quantum electrodynamics and quantum chromodynamics suggests a world with a small number of parameters and a unique ground state. For the most part, string theorists bought into this paradigm. At first it was hoped that string theory would be unique and explain the various parameters that quantum field theory left unexplained. When this turned out to be false, the belief developed that there were exactly five string theories with names like type–2a and Heterotic. This also turned out to be wrong. Instead, a continuum of theories were discovered that smoothly

---

One scenario is that, rather than quantum tunnelling, a particle accelerator, which produces very high energies in a very small area, could create sufficiently high energy density as to penetrate the barrier and stimulate the decay of the false vacuum to the lower energy vacuum. Hut and Rees [Hut and Rees (1983)], however, have determined that because we had observed cosmic ray collisions at much higher energies than those produced in terrestrial particle accelerators, that these experiments will not, at least for the foreseeable future, pose a threat to our vacuum. Particle accelerations have reached energies of only approximately four thousand billion electron volts (4 × 10^{33} GeV). Cosmic ray collisions have been observed at and beyond energies of 10^{11} GeV, the so–called Greisen–Zatsepin–Kuzmin limit. This event would be contingent on our living in a metastable vacuum, an issue which is far from resolved [Turner and Wilczek (1982)]. Worries about the vacuum metastability event are reminiscent of the controversy about turning the Relativistic Heavy Ion Collider on.
interpolated between the five and also included a theory called M–Theory. The language changed a little. One no longer spoke of different theories, but rather different solutions of some master theory. The space of these solutions is called the moduli space of supersymmetric vacua. Susskind calls it the supermoduli–space. Moving around on this supermoduli–space is accomplished by varying certain dynamical moduli. Examples of moduli are the size and shape parameters of the compact internal space that 4–dimensional string theory always needs. These moduli are not parameters in the theory but are more like fields. As you move around in ordinary space, the moduli can vary and have their own equations of motion. In a low energy approximation the moduli appear as massless scalar fields. The beauty of the supermoduli–space point of view is that there is only one theory but many solutions which are characterized by the values of the scalar field moduli. The mathematics of the string theory is so precise that it is hard to believe that there is not a consistent mathematical framework underlying the supermoduli–space vacua.

However the continuum of solutions in the supermoduli–space are all supersymmetric with exact super–particle degeneracy and vanishing cosmological constant. Furthermore they all have massless scalar particles, the moduli themselves. Obviously none of these vacua can possibly be our world. Therefore the string theorist must believe that there are other discrete islands lying off the coast of the supermoduli–space. The hope now is that a single non–supersymmetric island or at most a small number of islands exist and that non–supersymmetric physics will prove to be approximately unique. This view is not inconsistent with present knowledge (indeed it is possible that there are no such islands) but we find it completely implausible. It is much more likely that the number of discrete vacua is astronomical [Susskind (2003)].

This change in viewpoint is demanded by two facts, one observational and one theoretical. The first is that the expansion of the universe is accelerating. The simplest explanation is a small but non–zero cosmological constant. Evidently we have to expand our thinking about vacua to include states with non–zero vacuum energy. The incredible smallness and apparent fine tuning of the cosmological constant makes it absurdly improbable to find a vacuum in the observed range unless there are an enormous number of solutions with almost every possible value of $\lambda$. It seems to me inevitable that if we find one such vacuum we will find a huge number of them. We will from now on call the space of all such string theory vacua the landscape.

The second fact is that some recent progress has been made in exploring
the landscape \cite{Bousso and Polchinski (2000), Kachru et al. (2003)}. Before explaining the new ideas we need to define more completely what we mean by the landscape. The supermoduli–space is parameterized by the moduli which we can think of as a collection of scalar fields $\Phi_n$. Unlike the case of Goldstone bosons, points in the moduli space are not related by a symmetry of the theory. Generically, in a quantum field theory, changing the value of a non–Goldstone scalar involves a change of potential energy. In other words there is a non–zero field potential $V(\Phi)$. Local minima of $V$ are what we call vacua. If the local minimum is an absolute minimum the vacuum is stable. Otherwise it is only metastable. The value of the potential energy at the minimum is the cosmological constant for that vacuum.

To the extent that the low energy properties of string theory can be approximated by field theory, similar ideas apply. Bearing in mind that the low energy approximation may break down in some regions of the landscape, we will assume the existence of a set of fields and a potential. The space of these fields is the landscape.

The supermoduli–space is a special part of the landscape where the vacua are supersymmetric and the potential $V(\Phi)$ is exactly zero. These vacua are marginally stable and can be excited by giving the moduli arbitrarily small time derivatives. On the supermoduli–space the cosmological constant is also exactly zero. Roughly speaking, the supermoduli–space is a perfectly flat plain at exactly zero altitude (the value of $V$). Once we move off the plain, supersymmetry is broken and a non–zero potential develops, usually through some non–perturbative mechanism. Thus beyond the flat plain we encounter hills and valleys. We are particularly interested in the valleys where we find local minima of $V$. Each such minimum has its own vacuum energy. The typical value of the potential difference between neighboring valleys will be some fraction of $M_p^4$ where $M_p$ is the Planck mass. The potential barriers between minima will also be of similar height. Thus if a vacuum is found with cosmological constant of order $10^{-120}M_p^4$, it will be surrounded by much higher hills and other valleys.

Next consider two large regions of space, each of which has the scalars in some local minimum, the two minima being different. If the local minima are landscape–neighbors then the two regions of space will be separated by a domain wall. Inside the domain wall the scalars go over a “mountain pass”. The interior of the regions are vacuum like with cosmological constants. The domain wall which can also be called a membrane has additional energy in the form of a membrane tension. Thus there will be configurations of string
theory which are not globally described by a single vacuum but instead consists of many domains separated by domain walls. Accordingly, the landscape in field space is reflected in a complicated terrain in real space.

There are scalar fields that are not usually thought of as moduli but once we leave the flat plain there is no any fundamental difference. These are the four–form field strengths first introduced in the context of the cosmological constant by Brown and Teitelboim [Brown and Teitelboim (1988)]. A simple analogy exists to help visualize these fields and their potential. Think of 1+1 dimensional electrodynamics with electric fields \( E \) and massive electrons. The electric field is constant in any region of space where there are no charges. The field energy is proportional to the square of the field strength. The electric field jumps by a quantized unit whenever an electron is passed. Going in one direction, say along the positive \( x \) axis, the field makes a positive unit jump when an electron is passed and a negative jump when a positron is passed. In this model different vacua are represented by different quantized values of the electric field while the electrons/positrons are the domain walls. The energy of a vacuum is proportional to \( E^2 \). This model is not fundamentally different than the case with scalar fields and a potential. In fact by bosonizing the theory it can be expressed as a scalar field theory with a potential:

\[
V(\phi) = c\phi^2 + \mu \cos \phi.
\]

If \( \mu \) is not too small there are many minima representing the different possible 2–form field strengths, each with a different energy.

In 3 + 1 dimensions the corresponding construction requires a 4–form field strength \( F \) whose energy is also proportional to \( F^2 \). This energy appears in the gravitational field equations as a positive contribution to the cosmological constant. The analogue of the charged electrons are membranes which appear in string theory and function as domain walls to separate vacua with different \( F \). This theory can also be written in terms of a scalar field with a potential similar to \( V(\phi) \).

Let’s now consider a typical compactification of M–theory from eleven to 4 dimensions. The simplest example is gotten by choosing for the compact directions a 7–torus. The torus has a number of moduli representing the sizes and angles between the seven 1–cycles. The 4–form fields have as their origin a 7–form field strength, which is one of the fundamental fields of M–theory. The 7–form fields have 7 anti–symmetrized indecies. These non–vanishing 7–form can be configured so that three of the indecies are identified with compact dimensions and the remainder with uncompactified

---

\[20\]In 1 + 1 dimensions there is no magnetic field and the electric field is a two form, aka, a scalar density.
space–time. This can be done in thirty five \( = (7 \times 6 \times 5)/(1 \times 2 \times 3) \) ways which means that there are that many distinct 4–form fields in the uncompactified non–compact space. More generally, in the kinds of compact manifolds used in string theory to try to reproduce standard model physics there can be hundreds of independent ways of “wrapping” three compact directions with flux, thus producing hundreds of 3+1 dimensional 4–form fields. As in the case of 1+1 dimensional quantum electrodynamics, the field strengths are quantized, each in integer multiples of a basic field unit. A vacuum is specified by a set of integers \( n_1, n_2, ..., n_N \) where \( N \) can be as big as several hundred or more. The energy density of the energy of the 4–form fields has the form: \( \epsilon = c_i N_i^2 \), where the constants \( c_i \) depend on the details of the compact space.

The analogue of the electrons and positrons of the 1+1 dimensional example are branes. The 11 dimensional M–theory has 5–branes which fill 5 spatial directions and time. By wrapping 5–branes the same way the fluxes of the 4–forms are wrapped on internal 3–cycles leaves 2–dimensional membranes in 3+1 dimensions. These are the domain walls which separate different values of field strength. There are \( N \) types of domain wall, each allowing a unit jump of one of the 4–forms.

Bousso and Polchinski [Bousso and Polchinski (2000)] begin by assuming they have located some deep minimum of the field potential at some point \( \Phi_0 \). The value of the potential is supposed to be very negative at this point, corresponding to a negative cosmological constant , \( \lambda_0 \) of order the Planck scale. Also the 4–forms are assumed to vanish at this point. They then ask what kind of vacua can they get by discretely increasing the 4–forms. The answer depends to some degree on the compactification radii on the internal space but with modest parameters it is not hard to get such a huge number of vacua that it is statistically likely to have one in the range \( \lambda \sim 10^{-120} M_4^4 \).

To see how this works we write the cosmological constant as the sum of two terms, is the cosmological constant for vanishing 4–form, and the contribution of the 4–forms: \( \lambda = \lambda_0 + c_i N_i^2 \). With a hundred terms and modestly small values for the \( c_i \) it is highly likely to find a value of \( \lambda \) in the observed range. Note that no fine tuning is required, only a very large number of ways to make the vacuum energy.

The problem with [Bousso and Polchinski (2000)] was clearly recognized by the authors; The starting point is so far from the supermoduli–space that none of the usual tools of approximate supersymmetry are available to control the approximation. The example was intended only as a model
of what might happen because of the large number of possibilities.

More recently Kachru, Kallosh, Linde and Trivedi [Kachru et al. (2003)] have improved the situation by finding an example which is more under control. These authors subtly use the various ingredients of string theory including fluxes, branes, anti–branes and instantons to construct a rather tractable example with a small positive cosmological constant.

In addition to arguing that string theory does have many vacua with positive cosmological constant the argument in [Kachru et al. (2003)] tends to dispel the idea that vacua, not on the supermoduli–space , must have vanishing cosmological constant. In other words there is no evidence in string theory that a hoped for but unknown mechanism will automatically force the cosmological constant to zero. It seems very likely that all of the non–supersymmetric vacua have finite $\lambda$.

The vacua in [Kachru et al. (2003)] are not at all simple. They are jury–rigged, Rube Goldberg contraptions that could hardly have fundamental significance. But in an anthropic theory simplicity and elegance are not considerations. The only criteria for choosing a vacuum is utility, i.e., does it have the necessary elements such as galaxy formation and complex chemistry that are needed for life. That together with a cosmology that guarantees a high probability that at least one large patch of space will form with that vacuum structure is all we need [Susskind (2003)].

4.3.5.2 The Trouble with de Sitter Space

The classical vacuum solution of Einstein’s equations with a positive cosmological constant is de Sitter space. It is doubtful that it has a precise meaning in a quantum theory such as string theory [Goheer et al. (2002); Dyson et al. (2002b); Dyson et al. (2002a)]. We want to review some of the reasons for thinking that de Sitter space is at best a metastable state.

It is important to recognize that there are two very different ways to think about de Sitter space. The first is to take a global view of the space–time. The global geometry is described by the metric

$$ds^2 = R^2 \left\{ dt^2 - (\cosh t)^2 d^2 \Omega_3 \right\},$$

where $d^2 \Omega_d$ is the metric for a unit d–sphere and $R$ is related to the cosmological constant by

$$R = (\lambda G)^{-1/2}.$$

Viewing de Sitter space globally would make sense if it were a system
that could be studied from the outside by a ‘meta–observer’. Naively, the
meta–observer would make use of a (time dependent) Hamiltonian to evolve
the system from one time to another. An alternate description would use
a Wheeler de Witt formalism to define a wave function of the universe on
global space–like slices.

The other way of describing the space is the causal patch description.
The relevant metric is

$$ds^2 = R^2 \left\{ \left( 1 - r^2 \right) dt^2 - \frac{1}{\left( 1 - r^2 \right)} dr^2 - r^2 d^2 \Omega_2 \right\}.$$ 

In this form the metric is static and has a form similar to that of a black
hole. In fact the geometry has a horizon at $r = 1$. The static patch does
not cover the entire global de Sitter space but is analogous to the region
outside a black hole horizon. It is the region which can receive signals from,
and send signals to, an observer located at $r = 0$. To such an observer de
Sitter space appears to be a spherical cavity bounded by a horizon a finite
distance away [Susskind (2003)].

Experience with black holes has taught us to be very wary of global
descriptions when horizons are involved. In a black hole geometry there is
no global conventional quantum description of both sides of the horizon.
This suggests that a conventional quantum description of de Sitter space
only makes sense within a given observer’s causal patch. The descriptions in
different causal patches are complementary [Susskind et al. (1993);
Stephens et al. (1994)] but can not be put together into a global description
without somehow modifying the rules of quantum mechanics.

As in the black hole case, a horizon implies a thermal behavior with a
temperature and an entropy. These are given by: $T = 1/2\pi R$, $S = \pi R^2/G$.
We will assume the causal patch description of some particular observer.

If the observed ‘dark energy’ in the universe really is a small positive
cosmological constant the ultimate future of our universe will be eternal de
Sitter space. This would mean not that the future is totally empty space
but that the world will have all the features of an isolated finite thermal
avity with finite temperature and entropy. Thermal equilibrium for such
a system is not completely featureless. On short time scales not much can
be expected to happen but on very long time scales everything happens. A
famous example involves a gas of molecules in a sealed room. Imagine that
we start all the molecules in one corner of the room. In a relatively short
time the gas will spread out to fill the room and come to thermal equilib-
rium. During the approach to equilibrium interesting dissipative structures
such as droplets, eddies and vortices form and then dissipate. The usual assumption is that nothing happens after that. The entropy has reached its maximum value and the second law forbids any further interesting history. But on a sufficiently long time scale, large fluctuations will occur. In fact the phase point will return over and over to the neighborhood of any point in phase space including the original starting point. These Poincare recurrences generally occur on a time scale exponentially large in the thermal entropy of the system. Thus we define the Poincare recurrence time as: $T_r = e^S$. On such long time scale the second law of thermodynamics will repeatedly be violated by large scale fluctuations.

Thus even a pure de Sitter space would have an interesting cosmology of sorts. The causal patch of any observer would undergo Poincare recurrences in which it would endlessly fluctuate back to a state similar to its starting point, but each time slightly different.

The trouble with such a cosmology is that it relies on very rare “miracles” to start it off each time. But there are other miracles which could occur and lead to anthropically acceptable worlds with a vastly larger probability than our world. Roughly speaking the relative probability of a fluctuation leading to a given configuration is proportional to the exponential of its entropy. An example of a configuration far more likely than our own would be a world in which everything would be just like our universe except the temperature of the cosmic microwave background was ten degrees instead of three. When we say everything is the same we are including such details as the abundance of the elements.

Ordinarily such a universe would be ruled out on the grounds that it would take a huge miracle for the helium and deuterium to survive the bombardment by the extra photons implied by the higher temperature. That is correct, a fantastic miracle would be required, but such miracles would occur far more frequently than the ultimate miracle of returning to the starting point. This can be argued just from the fact that a universe at 10 degrees K has a good deal more entropy than one at 3 degrees. In a world based on recurrences it would be overwhelming unlikely that cosmology could be traced back to something like the inflationary era without a miraculous reversal of the second law along the way. Thus we are forced to conclude that the sealed tin can model of the universe must be incorrect, at least for time scales as long as the recurrence time.

Another difficulty with an eternal de Sitter space involves a mathematical conflict between the symmetry of de Sitter space and the finiteness of the entropy [Goheer et al. (2002)]. Basically the argument is that the finite-
ness of the de Sitter space entropy indicates that the spectrum of energy is discrete. It is possible to prove that the symmetry algebra of de Sitter space can not be realized in a way which is consistent with the discreteness of this spectrum. In fact this problem is not independent of the issues of recurrences. The discreteness of the spectrum means that there is a typical energy spacing of order: $\Delta E \sim e^{-S}$. The discreteness of the spectrum can only manifest itself on time scales of order $(\Delta E)^{-1}$ which is just the recurrence time. Thus there are problems with realizing the full symmetries of de Sitter space for times as long as $T_r$.

Finally another difficulty for eternal de Sitter space is that it does not fit at all well with string theory. Generally the only objects in string theory which are rigorously defined are S–Matrix elements. Such an S matrix can not exist in a thermal background. Part of the problem is again the recurrences which undermine the existence of asymptotic states. Unfortunately there are no known observables in de Sitter space which can substitute for S–matrix elements. The unavoidable implication of the issues is that eternal de Sitter space is an impossibility in a properly defined quantum theory of gravity [Susskind (2003)].

### 4.3.5.3 De Sitter Space is Unstable

In [Kachru et al. (2003)] a particular string theory vacuum with positive $\lambda$ was studied. One of the many interesting things that the authors found was that the vacuum is unstable with respect to tunnelling to other vacua. In particular the vacuum can tunnel back to the supermoduli–space with vanishing cosmological constant. Using instanton methods the authors calculated that the lifetime of the vacuum is less than the Poincare recurrence time. This is no accident. To see why it always must be so, let’s consider the effective potential that the authors of [Kachru et al. (2003)] derived. The only modulus which is relevant is the overall size of the compact manifold $\Phi$. The de Sitter vacuum occurs at the point $\Phi = \Phi_0$. However, the absolute minimum of the potential occurs not at $\Phi_0$ but at $\Phi = \infty$. At this point the vacuum energy is exactly zero and the vacuum one of the ten dimensional vacua of the supermoduli–space. There are always runaway solutions like this in string theory. The potential on the supermoduli–space is zero and so it is always possible to lower the energy by tunnelling to a point on the supermoduli–space.

Suppose we are stuck in the potential well at $\Phi_0$. The vacuum of the causal patch has a finite entropy and fluctuates up and down the walls
of the potential. One might think that fluctuations up the sides of the potential are Boltzmann suppressed. In a usual thermal system there are two things that suppress fluctuations. The first is the Boltzmann suppression by factor $e^{-\beta E}$, and the second is entropy suppression by factor $e^{(S_f-S)}$, where $S$ is the thermal entropy and $S_f$ is the entropy characterizing the fluctuation which is generally smaller than $S$. However in a gravitational theory in which space is bounded (as in the static patch) the total energy is always zero, at least classically. Hence the only suppression is entropic. The phase point wanders around in phase space spending a time in each region proportional to its phase space volume, i.e., $e^{-S_f}$. Furthermore, the typical time scale for such a fluctuation to take place is of order $T_f \sim e^{(S-S_f)}$ [Susskind (2003)].

Now consider a fluctuation which brings the field $\phi$ to the top of the local maximum at $\phi = \phi_1$ in the entire causal patch. The entropy at the top of the potential is given in terms of the cosmological constant at the top. It is obviously positive and less than the entropy at $\phi_0$. Thus the time for the field to fluctuate to $\phi_1$ (over the whole causal patch) is strictly less than the recurrence time $e^S$. But once the field gets to the top there is no obstruction to it rolling down the other side to infinity. It follows that a de Sitter vacuum of string theory is never longer lived than $T_r$ and furthermore we end up at a supersymmetric point of vanishing cosmological constant.

There are other possibilities. If the cosmological constant is not very small it may tunnel over the nearest mountain pass to a neighboring valley of smaller positive cosmological constant. This will also take place on a time scale which is too short to allow recurrences. By the same argument it will not stay in the new vacuum indefinitely. It may find a vacuum with yet smaller cosmological constant to tunnel to. Eventually it will have to make a transition out of the space of vacua with positive cosmological constants.

4.3.5.4 Bubble Cosmology

To make use of the enormous diversity of environments that string theory is likely to bring with it, we need a dynamical cosmology which, with high
probability, will populate one or more regions of space with an anthropically favorable vacuum. There is a natural candidate for such a cosmology that we will explain from the global perspective (Susskind (2003)).

For simplicity let’s temporarily assume that there are only two vacua, one with positive cosmological constant \( \lambda \), and one with vanishing cosmological constant. Without worrying how it happened we suppose that some region of the universe has fallen into the minimum with positive cosmological constant. From the global perspective it is inflating and new Hubble volumes are constantly being produced by the expansion. Pick a time-like observer who looks around and sees a static universe bounded by a horizon. The observer will eventually observe a transition in which his entire observable region slides over the mountain pass and settles to the region of vanishing \( \lambda \). The observer sees the horizon–boundary quickly recede, leaving in its wake an infinite open Freedman, Robertson, Walker universe with negative spatial curvature. The final geometry has light–like and time like future infinities similar to flat space.

Now let us take the more global view. The bubble does not swallow the entire global space but leaves part of the space still inflating. Inevitably bubbles will form in this region. In fact if we follow the world line of any observer, it will eventually be swallowed by a bubble of \( \lambda = 0 \) vacuum.

The real landscape is not comprised of only two vacua. If an observer starts with a large value of the cosmological constant there will be many ways for the causal patch to descend to the supermoduli–space. From the global viewpoint a bubbles will form in neighboring valleys with somewhat smaller cosmological constant. Since each bubble has a positive cosmological constant it will be inflating but the space between bubbles is inflating faster so the bubbles go out of causal contact with one another. Each bubble evolves in isolation from all the others. Furthermore, in a time too short for recurrences, bubbles will nucleate within the bubbles. Following a single observer within his own causal patch, the cosmological constant decreases in a series of events until the causal patch finds itself in the supermoduli–space. Each observer will see a series of vacuums descending down to the supermoduli–space and the chances that he passes through an anthropically acceptable vacuum is most likely very small. But on the other hand the global space contains an infinite number of such histories and some of them will be acceptable.

The only problem with the cosmology that we just outlined is that it is formulated in global coordinates. From the viewpoint of any causal patch, all but one of the bubbles is outside the horizon. As I’ve emphasized, the
application of the ordinary rules of quantum mechanics only makes sense within the horizon of an observer. We do not know the rules for putting together the various patches into one comprehensive global description and until we do there cannot be any firm basis for the kind of anthropic cosmology I described. Nevertheless the picture is tempting [Susskind (2003)].

4.3.5.5 Cosmology as a Resonance

The idea of scalar fields and potentials is approximate once we leave the supermoduli–space. So is the notion of a stable de Sitter vacuum. The problem is familiar. How do we make precise sense of an unstable state in quantum mechanics. In ordinary quantum mechanics the clearest situation is when we can think of the unstable state as a resonance in a set of scattering amplitudes. The parameters of a resonance, i.e., its width and mass are well defined and do not depend on the exact way the resonance was formed. Thus even black holes have precise meaning as resonant poles in the S–matrix. Normally we cannot compute the scattering amplitudes that describe the formation and evaporation of a black hole but it is comforting that an exact criterion exists.

In the case of a black hole the density of levels is enormous being proportional to the exponential of the entropy. The spacing between levels is therefore exponentially small. On the other hand the width of each level is not very small. The lifetime of a state is the time that it takes to emit a single quantum of radiation and this is proportional to the Schwarzschild radius. Therefore the levels are broadened by much more than their spacing. The usual resonance formulas are not applicable but the precise definition of the unstable state as a pole in the scattering amplitude is. We think the same things can be said about the unstable de Sitter vacua but it can only be understood by returning to the causal patch way of thinking [Susskind (2003)]. Therefore let’s focus on the causal patch of one observer. We have discussed the observer’s future history and found that it always ends in an infinite expanding supersymmetric open Freedman universe. Such a universe has the usual kind of asymptotic future consisting of time–like and light–like infinity. There is no temperature in the remote future and the geometry permits particles to separate and propagate as free particles just as in flat space–time.

Now let’s consider the observer’s past history. The same argument which says that the observer will eventually make a transition to \( \lambda = 0 \) in the far future can be run backward. The observer could only have gotten to the
de Sitter vacuum by the time–reversed history so he must have originated from a collapsing open universe. The history may seem paradoxical since it requires the second law of thermodynamics to be violated in the past. A similar paradox arises in a more familiar setting. Let me return to the sealed room filled with gas molecules except that now one of the walls has a small hole that lets the gas escape to unbounded space. Suppose we find the gas filling the room in thermal equilibrium at some time. If we run the system forward we will eventually find all the molecules have escaped and are on their way out, never to return. But it is also true that if we run the equations of motion backwards we will eventually find all the molecules outside the room moving away. Thus the only way the starting configuration could have occurred is if the original molecules were converging from infinity toward the small hole in the wall.

If we are studying the system quantum mechanically, the metastable configuration with all the molecules in the room would be an unstable resonance in a scattering matrix describing the many body scattering of a system of molecules with the walls of the room. Indeed the energy levels describing the molecules trapped inside the room are complex due to the finite lifetime of the configuration.

This suggests a view of the intermediate de Sitter space as an unstable resonance in the scattering matrix connecting states in the asymptotic \( \lambda = 0 \) vacua. In fact we can estimate the width of the states. Since the lifetime of the de Sitter space is always longer than the recurrence time, generally by a huge factor, the width \( \gamma \) satisfies: \( \gamma \gg e^{-S} \). On the other hand the spacing between levels, \( \Delta E \), is of order \( e^{-S} \). Therefore, \( \gamma \gg \Delta E \), so that the levels are very broad and overlapping as for the black hole [Susskind (2003)].

No perfectly precise definition exists in string theory for the moduli fields or their potential when we go away from the supermoduli–space. The only precise definition of the de Sitter vacua seems to be as complex poles in some new sector of the scattering matrix between states on the supermoduli–space.

Knowing that a black hole is a resonance in a scattering amplitude does not tell us much about the way real black holes form. Most of the possibilities for black hole formation are just the time reverse of the ways that it evaporate. In other words the overwhelming number of initial states that can lead to a black hole consist of thermal radiation. Real black holes in our universe form from stellar collapse which is just one channel in a huge collection of S–matrix ‘in states’. In the same way the fact that
cosmological states may be thought of in a scattering framework is in itself
does not shed much light on the original creation process.

As seen above, vacua come in two varieties, supersymmetric and other-
wise. Most likely the non-supersymmetric vacua do not have vanishing
cosmological constant but it is plausible that there are so many of them
that they practically form a continuum. Some tiny fraction have cosmol-
ogical constant in the observed range. With nothing preferring one vacuum
over another, the anthropic principle comes to the fore whether or not we
like the idea. String theory provides a framework in which this can be
studied in a rigorous way. Progress can certainly be made in exploring the
landscape. The project is in its infancy but in time we should know just
how rich it is. We can argue the philosophical merits of the anthropic prin-
ciple but we can’t argue with quantitative information about the number
of vacua with each particular property such as the cosmological constant,
Higgs mass or fine structure constant. That information is there for us to
extract [Susskind (2003)].

Counting the vacua is important but not sufficient. More understanding
of cosmological evolution is essential to determining if the large number of
possibilities are realized as actualities. The vacua in string theory with \( \lambda > 0 \) are not stable and decay on a time scale smaller than the recurrence time.
This is very general and also very fortunate since there are serious problems
with stable de Sitter space. The instability also allows the universe to
sample all or a large part of the landscape by means of bubble formation.
In such a world the probability that some region of space has suitable
conditions for life to exist can be large.

The bubble universe based on Linde’s *eternal inflation* seems promising
but it is unclear how to think about it with precision. There are real concept-
tual problems having to do with the global view of space–time. The main
problem is to reconcile two pictures; the causal patch picture and the global
picture. String theory has provided a testing ground for some important
relevant ideas such as black hole complementarity [Susskind *et al.* (1993),
Stephens *et al.* (1994)] and the Holographic principle [‘t Hooft (2093);
Susskind (1995)]. Complementarity requires the observer’s side of the hori-
zon to have a self contained conventional quantum description. It also
prohibits a conventional quantum description that covers the interior and
exterior simultaneously. Any attempt to describe both sides as a single
quantum system will come into conflict with one of three sacred principles
[Susskind (2002)]. The first is the equivalence principle which says that a
freely falling observer passes the horizon without incident. The second says
that experiments performed outside a black hole should be consistent with the rules of quantum mechanics as set down by Dirac in his textbook. No loss of quantum information should take place and the time evolution should be unitary. Finally the rules of quantum mechanics forbid information duplication. This means that we can not resolve the so called information paradox by creating two copies (quantum Xeroxing) of every bit as it falls through the horizon; at least not within the formalism of conventional quantum mechanics. The complementarity and holographic principles have been convincingly confirmed by the modern methods of string theory [Maldacena (2001)]. The inevitable conclusion is that a global description of geometries with horizons, if it exists at all, will not be based on the standard quantum rules.

Why is this important for cosmology? The point is that the eternal inflationary production of an infinity of bubbles takes place behind the horizon of any given observer. It is not something that has a description within one causal patch. If it makes sense, a global description is needed but if cosmic event horizons are at all like black hole horizons then any global description will involve wholly new elements. If we were to make a wild guess about which rule of quantum mechanics has to be given up in a global description of either black holes or cosmology we would guess it is the Quantum Xerox Principle [Susskind (2002)]. We would look for a theory which formally allowed quantum duplication but cleverly prevents any observer from witnessing it. Perhaps then the replication of bubbles can be sensibly described.

Progress may also be possible in sharpening the exact mathematical meaning of the de Sitter vacua. Away from the supermoduli–space, the concept of a local field and the effective potential is at best approximate in string theory. The fact that the vacua are false metastable states makes it even more problematic to be precise. In ordinary quantum mechanics the best mathematical definition of an unstable state is as a resonance is amplitudes for scattering between very precisely defined asymptotic states. Each metastable state corresponds to pole whose real and imaginary parts define the energy and inverse lifetime of the state.

We have argued that each causal patch begins and ends with an asymptotic ‘roll’ toward the supermoduli–space [Susskind (2003)]. The final state have the boundary conditions of an FRW open universe and the initial states are time reversals of these. This means we may be able to define some kind S–matrix connecting initial and final asymptotic states. The various intermediate metastable de Sitter phases would be exactly defined.
as resonant resonances in this amplitude.

At first this proposal sounds foolish. In general relativity initial and final states are very different. Black holes make sense. White holes do not. Ordinary things fall into black holes and thermal radiation comes out. The opposite never happens. But this is deceiving. Our experience with string theory has made it clear that the fundamental micro–physical input is completely reversible and that black holes are most rigorously defined in terms of resonances in scattering amplitudes\(^ {22} \). Knowing that a black hole is an intermediate state in a tremendously complicated scattering amplitude does not really tell us much about how real black holes form. For that we need to know about stellar collapse and the like. But it does provide an exact mathematical definition of the states that comprise the black hole ensemble.

From the causal patch viewpoint the evolutionary endpoints seems to be an approach to some point on the supermoduli–space. After the last tunnelling the universe enters an final open FRW expansion toward some flat supersymmetric solution. This is not to be thought of as a unique quantum state but as a large set of states with similar evolution. Running the argument backward (assuming microscopic reversibility) we expect the initial state to be the time reversal of one of the many future endpoints. We might even hope for a scattering matrix connecting initial and final states. de Sitter minima would be an enormously large density of complex poles in the amplitude. For more details, see [Susskind (2003)].

### 4.3.6 Top–Down Cosmology

In this section, following [Hawking and Hertog (2002)], we try to convince the readers that inflation actually starts at the ‘top of the hill’.

Structure and complexity have developed in our universe, because it is out of equilibrium. This feature shows up in all known cosmological scenarios for the early universe, which rely on gravitational instability to generate local inhomogeneities from an almost homogeneous and isotropic state for the universe. Inflation seems the best explanation for this homogeneous and isotropic state because whatever drives the inflation will remove the local instability and iron out irregularities. However the inflationary expansion has to be globally unstable because otherwise it would continue forever and galaxies would never form.

The instability can be described as the evolution of an order parameter\(^ {22} \)The one exception is black holes in anti–de Sitter space which are stable.
\( \phi \) which can be treated as a scalar field with effective potential \( V(\phi) \). If \( V'/V \) is small, \( \phi \) will roll slowly down the potential and the universe will inflate by a large factor. However, this raises the question: Why did the universe start with a high value of the potential? Why didn’t \( \phi \) start at the global minimum of \( V \)?

There have been various attempts to explain why \( \phi \) started high on the potential hill. In the old [Guth (1982)] and new [Linde (1982); Albrecht and Steinhardt (1982)] inflationary scenarios the universe was supposed to start with infinite temperature at a singularity. As the universe expanded and cooled, thermal corrections would make the effective potential time dependent. So even if \( \phi \) started in the minimum of \( V \), it could still end up in a metastable false vacuum state (in old inflation) or at a local maximum of \( V \) (in new inflation). The scalar field was then supposed to tunnel through the potential barrier or just fall off the top of the hill and slowly roll down. However both scenarios tended to predict a more inhomogeneous universe than we observe. They were also unsatisfactory because they assumed an initial singularity and a fairly homogeneous and isotropic pre–inflation hot Big–Bang phase. Why not just assume the singularity produced the standard hot Big–Bang, since we don’t have a measure on the space of singular initial conditions for the universe?

In the chaotic inflation scenario [Linde (1983)], quantum fluctuations of \( \phi \) are supposed to drive the volume weighted average \( \phi \) up the potential hill, leading to everlasting eternal inflation. However this effect is dependent on using the synchronous gauge: in other gauges the volume weighted average of the potential can go down. Looking from a 4 rather than 3+1 dimensional perspective, it is clear that the quantum fluctuations of a single scalar field are insufficient to drive de Sitter like eternal inflation, if the de Sitter space is larger than the Planck length. Eternal inflation may be possible at the Planck scale, but all presented methods would break down in this situation so it would mean that we could not analyze the origin of the universe.

The aim of this section however is to show that the universe can come into being and start inflating without the need for an initial hot Big–Bang phase or Planck curvature. It is required that the potential \( V \) has a local maximum which is below the Planck density and sufficiently flat on top, \( V''/V > -4/3 \). This last condition means only the homogeneous mode of the scalar field is tachyonic: the higher modes all have positive eigenvalues. It also means there is not a Coleman–De Luccia solution [Coleman (1980)] describing quantum tunnelling from a false vacuum on one side of the maximum to the true vacuum on the other side. Instead there is only
a homogeneous Hawking–Moss instanton [Hawking and Moss (1982)] that sits on the top of the hill, at the local maximum of $V$.

It has long been a problem to understand how the universe could decay from a false vacuum in this situation. The Hawking–Moss instanton does not interpolate between the false and true vacua, because it is constant in space and time. Instead, what must happen is that the original universe can continue in the false vacuum state but that new completely disconnected universes can form at the top of the hill via Hawking–Moss instantons. For someone in one of these new universes, the universe in the false vacuum is irrelevant and can be ignored.

The top of the hill might seem the least likely place for the universe to start. However we shall show it is the most likely place for an inflationary universe to begin, if $V''/V > -4/3$. The reason is that although being at the top of the hill costs potential action, the saving of gradient action from having a constant scalar field is greater. Thus inflation will start at the top of the hill. In particular, this justifies Starobinsky’s scenario of trace anomaly inflation, in which the universe starts in an unstable de Sitter state supported by the conformal anomaly of a large number of conformally coupled matter fields [Starobinsky (1980)].

The usual approach to the problem of initial conditions for inflation, is to assume some initial configuration for the universe, and evolve it forward in time. This could be described as the bottom up approach to cosmology. It is an essentially classical picture, because it assumes there is a single well defined metric for the universe. By contrast, here we adopt a quantum approach, based on the no boundary proposal [Hartle and Hawking (1983)], which states that the amplitude for an observable like the 3–metric on a space–like hypersurface $\Sigma$, is given by a path integral over all metrics whose only boundary is $\Sigma$. The quantum origin of our universe and the no boundary proposal naturally lead to a top down view of the universe, in which the histories that contribute to the path integral, depend on the observable being measured.

Following [Hawking and Hertog (2002)], we study the quantum cosmological origin of an expanding universe in theories like trace anomaly inflation, by investigating the semiclassical predictions of the no boundary proposal for the wave function of interest. One may argue that a clearer picture of the pre–inflationary conditions can only emerge from a deeper understanding of quantum gravity at the Planck scale. However, the amplitude of the cosmic microwave temperature anisotropies indicates that the
universe may always have been much larger than the Planck scale. This suggests it might be possible to describe the origin of our universe within the semiclassical regime of quantum cosmology. Correspondingly, the effective potential must have a local maximum well below the Planck density, which is the case in the trace anomaly model.

4.3.6.1 Trace Anomaly Driven Inflation

Large $N$ Cosmology

It has been argued that the theoretical foundations for inflation are weak, since it has proven difficult to realise inflation in classical M–theory. A large class of supergravity theories admit no warped de Sitter compactifications on a compact, static internal space \cite{Gibbons1985, Maldacena2001} and although some gauged $N = 8$ and $N = 4$ supergravities in $D = 4$ do permit de Sitter vacua \cite{Gates1983, Hull1984}, these vacua are too unstable for a significant period of inflation to occur. However, an appealing way to evade the no go theorems is to include higher derivative quantum corrections to the classical supergravity equations, such as the trace anomaly.

Since we observe a large number of matter fields in the universe, it is natural to consider the large $N$ approximation \cite{Tomboulis1977}. In the large $N$ approximation, one performs the path integral over the matter fields in a given background to get an effective action that is a functional of the background metric \cite{Hawking2002}.

$$\exp(-W[g]) = \int \mathcal{D}[\phi] \exp(-S[\phi, g]).$$

In the leading–order $1/N$ approximation, one can neglect graviton loops and look for a stationary point of the effective action for the matter fields combined with the gravitational action. This is equivalent to solving the Einstein equations with the source being the expectation value of the matter energy–momentum tensor derived from $W$,

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G\langle T_{\mu\nu} \rangle.$$

The expectation value of the energy–momentum tensor is generally non–local and depends on the quantum state. However, during inflation, particle masses are small compared with the space–time curvature, $R \gg m^2$, and in asymptotically free gauge theories, interactions become negligible in the
same limit. Therefore, at the high curvatures during inflation, the energy–momentum tensor of a large class of grand unified theories is to a good approximation given by the expectation value \( \langle T_{\mu\nu} \rangle \) of a large number of free, massless, conformally invariant fields.\(^{23}\) The entire one-loop contribution to the trace of the energy–momentum tensor then comes from the conformal anomaly \([\text{Capper (1974)}]\), which is given for a general CFT by the following equation,

\[
g^{\mu\nu} \langle T_{\mu\nu} \rangle = cF - aG + \alpha' \nabla^2 R,
\]  

(4.77)

where \( F \) is the square of the Weyl tensor, \( G \) is proportional to the Euler density and the constants \( a, c \) and \( \alpha' \) are given in terms of the field content of the CFT by

\[
a = \frac{1}{360 (4\pi)^2} (N_S + 11N_F + 62N_V), \quad c = \frac{1}{120 (4\pi)^2} (N_S + 6N_F + 12N_V),
\]

\[
\alpha' = \frac{1}{180 (4\pi)^2} (N_S + 6N_F - 18N_V),
\]

with \( N_S \) the number of real scalar fields, \( N_F \) the number of Dirac fermions and \( N_V \) the number of vector fields.

The trace anomaly is entirely geometrical in origin and therefore independent of the quantum state. In a maximally symmetric space–time, the symmetry of the vacuum implies that the expectation value of the energy–momentum tensor is proportional to the metric,

\[
\langle 0 | T_{\mu\nu} | 0 \rangle = \frac{1}{4} g_{\mu\nu} g^{|\rho\sigma\rangle} \langle 0 | T_{\rho\sigma} | 0 \rangle.
\]

Thus the trace anomaly acts just like a cosmological constant for these space–times, and a positive trace anomaly permits a de Sitter solution to the Einstein equations.

The radius of the de Sitter solution is determined by the number of fields, \( N^2 \), in the CFT and is of order \( \sim Nl_{pl} \). Therefore the one-loop contributions to the energy–momentum tensor are \( \sim 1/N^2 \), which means they are of the same order as the classical terms in the Einstein equations. On the other hand, the corrections due to graviton loops are \( \sim 1/N^3 \), so for large \( N \) quantum gravitational fluctuations are suppressed, confirming the consistency of the large \( N \) approximation.

\(^{23}\)For simplicity, it is assumed that scalar fields become conformally coupled at high energies, but the contribution of the interaction terms to \( \langle T_{\mu\nu} \rangle \) is small at high curvature, as long as the couplings don’t become very large \([\text{Parker and Toms (1984)}]\).
For $\alpha' = 0$ in (4.77), the only $O(3, 1)$ invariant solutions are de Sitter space and flat space, which are the initial and final stages of the simplest inflationary universe. In order for a solution to exist that interpolates between these two stages, one must have $\alpha' < 0$ in (4.77), as Starobinsky discovered [Starobinsky (1980)]. Starobinsky showed that if $\alpha' < 0$, the de Sitter solution is unstable, and decays into a matter dominated Friedman–Lemaître Robertson–Walker universe, on a timescale determined by $\alpha'$. The purpose of Starobinsky’s work was to demonstrate that quantum effects of matter fields might resolve the Big–Bang singularity. From a modern perspective, it is more interesting that the conformal anomaly might have been the source of a finite, but significant period of inflation in the early universe. Rapid oscillations in the expansion rate at the end of inflation, would result in particle production and (p)reheating.

Starobinsky showed that the de Sitter solution is unstable both to the future and to the past, so it was not clear how the universe could have entered the de Sitter phase. This is the problem of initial conditions for trace anomaly driven inflation, which should be addressed within the framework of quantum cosmology, by combining inflation with a theory for the wave $\Psi$—function of the quantum universe. Hartle and Hawking suggested that the amplitude for the quantum state of the universe described by 3–metric $h$ and matter fields $\phi(x)$ on a 3–surface $\Sigma$, should be given by [Hartle and Hawking (1983)]

$$\Psi[\Sigma, h, \phi_\Sigma] = N \sum_M \int \mathcal{D}[g] \mathcal{D}[\phi(x)] e^{-S_E(g, \phi)},$$

(4.78)

where the Euclidean path integral is taken over all compact four geometries bounded only by a 3–surface $\Sigma$, with induced metric $h$ and matter fields $\phi_\Sigma$. $M$ denotes a diffeomorphism class of 4–manifolds and $N$ is a normalization factor. The motivation to restrict the class of manifolds and metrics to geometries with only a single boundary is that in cosmology, in contrast with scattering calculations, one is interested in measurements in a finite region in the interior of space–time. The ‘no boundary’ proposal gives a definite ansatz for the wave function $\Psi[\Sigma, h, \phi_\Sigma]$ of the universe and in principle removes the initial singularity in the hot Big–Bang model. At least within the semiclassical regime, this yields a well–defined probability measure on the space of initial conditions for cosmology.

One can appeal to quantum cosmology to explain how the de Sitter phase emerges in trace anomaly inflation, since the no boundary proposal can describe the creation of an inflationary universe from nothing. At the
semiclassical level, this process is mediated by a compact instanton saddle-point of the Euclidean path integral, which extrapolates to a real Lorentzian universe at late times. To find the relative probability of different geometries in the no boundary path integral, one must compute their Euclidean action. Below we consider a model of anomaly–induced inflation consisting of gravity coupled to $\mathcal{N} = 4$, $U(N)$ super Yang–Mills theory, for which the AdS/CFT correspondence [Maldacena (1998)] provides an attractive way to calculate the effective matter action on backgrounds without symmetry. The fact that we are using $\mathcal{N} = 4$, $U(N)$ super Yang–Mills theory is probably not significant since, as we shall describe, it is the large number of fields that matters here and not the Yang–Mills coupling. Therefore, we expect the presented results to be valid for any matter theory that is approximately massless during the de Sitter phase [Hawking and Hertog (2002)].

**Effective Matter Action**

We consider, in Euclidean signature, Einstein gravity coupled to a $\mathcal{N} = 4$, $U(N)$ super Yang–Mills theory with large $N$,

$$ S = -\frac{1}{2\kappa} \int d^4x\sqrt{g}R - \frac{1}{\kappa} \int d^3x\sqrt{h}K + W, \quad (4.79) $$

where $W$ denotes the Yang–Mills effective action. The field content of the Yang–Mills theory is $N_S = 6N^2$, $N_F = 2N^2$ and $N_V = N^2$, yielding an anomalous trace [Hawking and Hertog (2002)]

$$ g^{\mu\nu}(T_{\mu\nu}) = \frac{N^2}{64\pi^2}(F - G). $$

The one–loop result for the conformal anomaly is exact, since it is protected by supersymmetry. Therefore, inflation supported by the trace anomaly of $\mathcal{N} = 4$, $U(N)$ super Yang–Mills would never end. The presence of non–conformally invariant fields in realistic matter theories, however, necessarily alters the value of $\alpha'$ in the anomaly (4.77). Since the coefficient of the $\nabla^2 R$ term plays such an important role in trace anomaly driven inflation, we ought to include this correction. As a first approximation, one can account for the non–conformally invariant fields by adding a local counterterm to the action,

$$ S_{ct} = \frac{\alpha N^2}{192\pi^2} \int d^4x\sqrt{g}R^2. $$
This leads to an extra contribution to the conformal anomaly, which becomes
\[ g^{\mu\nu} \langle T_{\mu\nu} \rangle = \frac{N^2}{64\pi^2} (F - G) + \frac{\alpha N^2}{16\pi^2} \nabla^2 R. \]

For \( \alpha < 0 \), the expansion now changes from exponential to the typical power law \( \sim t^{2/3} \) of a matter dominated universe, on a time scale \( \sim 12|\alpha| \log N \).

One can construct more sophisticated models of anomaly driven inflation, by taking in account corrections from particle masses and interactions in a more precise way. One could, for instance, consider soft supersymmetry breaking during inflation. The coefficient \( \alpha' \) could then vary in time, because the decoupling of massive particles at low energy [Shapiro and Sola (2001)] alters the number of degrees of freedom that contribute to the quantum effective action. For our purposes, however, it is sufficient to consider the theory above.

In no boundary cosmology, one is interested in solutions that describe a Lorentzian inflationary universe that emerges from a compact instanton solution of the Euclidean field equations. These geometries provide saddle-points of the Euclidean path integral (4.98) for the wave function of interest. Because our universe is Lorentzian at late times, it has been suggested that the relevant instanton saddle-points of the no boundary path integral are so-called ‘real tunnelling’ geometries [Gibbons and Hartle (1990); Halliwell and Hartle (1990)]. Cosmological real tunnelling solutions are compact Riemannian geometries joined to an \( O(3,1) \) invariant Lorentzian solution of Einstein’s equations, across a hypersurface of vanishing extrinsic curvature \( K_{\mu\nu} \). Such instanton solutions can then be used as background in a perturbative evaluation of the no boundary path integral, to find correlators of metric perturbations during inflation, which in turn determine the cosmic microwave anisotropies.

We now compute the effective matter action \( W \) on such perturbed instanton metrics. After eliminating the gauge freedom, the perturbed metric on the spaces of interest can be written as
\[ ds^2 = B^2(\chi)\gamma_{\mu\nu} dx^\mu dx^\nu = B^2(\chi)((1 + \psi)\hat{\gamma}_{\mu\nu} + \theta_{\mu\nu}) dx^\mu dx^\nu, \]

where \( \hat{\gamma}_{\mu\nu} \) is the metric on the unit \( S^4 \) and \( \theta_{\mu\nu} \) is transverse and traceless with respect to the four sphere.

In order to evaluate the no boundary path integral, we must first compute the quantum effective action \( W[B, \hbar] \) on the background (4.80). The effective action of the matter fields is computed as an expansion around the
homogeneous background with metric
\[ g_{\mu\nu} = B^2(\chi) \hat{\gamma}_{\mu\nu}. \]

To second order in the metric perturbation, \( W[B, h] \) is determined by the one and 2–point function of the energy–momentum tensor on the unperturbed \( O(4) \) invariant background. The one-point function is given by the conformal anomaly. Since the FLRW background is conformal to the round four sphere, the 2–point function can be calculated by a conformal transformation from \( S^4 \). On \( S^4 \), the 2–point function is determined entirely by symmetry and the trace anomaly [Hawking et al. (2001)]. Therefore, since the energy–momentum tensor transforms anomalously, the 2–point function on (4.80) should be fully determined by the 2–point function on \( S^4 \), the trace anomaly and the scale factor \( B(\chi) \). For the matter theory we have in mind, all these quantities are independent of the coupling, so it follows that the effective action \( W[B, h] \) is independent of the coupling, to second order in the metric perturbation.

In [Riegert (1984)], it was found how the effective action that generates a conformal anomaly of the form (4.77), transforms under a conformal transformation. We can use this result to relate \( W[B, h] \) on the perturbed FLRW space to \( W[r, h] \) on the perturbed four sphere with radius \( r \). Writing \( B(\chi) = r e^{\sigma(\chi)} \), where \( r \) is an arbitrary radius, the transformation is given by [Hawking and Hertog (2002)]

\[
W[\sigma(\chi), h] = \tilde{W}[r, h] - \frac{N^2}{32\pi^2} \int d^4x \sqrt{\gamma} \left[ \sigma \left( R^{\mu\nu} R_{\mu\nu} - \frac{1}{3} R^2 \right) + 2 \nabla_{\mu} \sigma \nabla^\mu \sigma \nabla^2 \sigma + 2 \left( R^{\mu\nu} - \frac{1}{2} \gamma^{\mu\nu} R \right) \nabla_{\mu} \sigma \nabla_{\nu} \sigma + (\nabla_{\mu} \sigma \nabla^\mu \sigma)^2 \right].
\]

Here \( \tilde{W} \) denotes the effective action on the perturbed four sphere of radius \( r \) with metric \( \gamma_{\mu\nu} \), and the Ricci scalar \( R \) and covariant derivative \( \nabla_\mu \) refer to the same space.

The generating functional \( \tilde{W}[r, h] \) was computed in [Hawking et al. (2001)], by using the AdS/CFT correspondence [Maldacena (1998)]

\[
Z[h] \equiv \int \mathcal{D}[g] \exp(-S_{grav}[g]) = \int \mathcal{D}[\phi] \exp(-S_{CFT}[\phi; h]) \equiv \exp(-W_{CFT}[h]),
\]

where \( Z[h] \) is the supergravity partition function on \( AdS_5 \). The AdS/CFT calculation is performed by introducing a fictional ball of (Euclidean) AdS...
that has the perturbed sphere as its boundary. In the classical gravity limit, the CFT generating functional can then be obtained by solving the IIB supergravity field equations, to find the bulk metric $g$ that matches onto the boundary metric $h$, and adding a number of counterterms that depend on the geometry of the boundary, in order to render the action finite as the boundary is moved off to infinity. To second order in the perturbation $h$, the quantum effective action (including the $R^2$ counterterm) is given by

$$\tilde{W} = \tilde{W}^{(0)} + \tilde{W}^{(1)} + \tilde{W}^{(2)} + \ldots$$

where

$$\tilde{W}^{(0)} = -\frac{3\beta N^2 \Omega_4}{8\pi^2} + \frac{3\alpha N^2 \Omega_4}{4\pi^2} + \frac{3N^2 \Omega_4}{32\pi^2} (4\log 2 - 1),$$

$$\tilde{W}^{(1)} = \frac{3N^2}{16\pi^2 r^2} \int d^4x \sqrt{\hat{\gamma}} \psi,$$

$$\tilde{W}^{(2)} = -\frac{3N^2}{64\pi^2 r^4} \int d^4x \sqrt{\hat{\gamma}} \left[ \psi \left( \hat{\nabla}^2 + 2 \right) \psi - \alpha \psi \left( \hat{\nabla}^4 + 4\hat{\nabla}^2 \right) \psi \right]$$

$$+ \frac{N^2}{256\pi^2 r^4} \sum_p \left( \int d^4x' \sqrt{\hat{\gamma}} H_{\mu\nu}^{(p)}(x') \right)^2 (\Psi(p) - 4\alpha p(p + 3)), $$

where $p$ labels the eigenvalues of the Laplacian $\hat{\nabla}^2$ on the round four sphere and

$$\Psi(p) = p(p + 1)(p + 2)(p + 3) [\psi(p/2 + 5/2) + \psi(p/2 + 2) - \psi(2) - \psi(1)]$$

$$+ p^4 + 2p^3 - 5p^2 - 10p - 6 + 2\beta p(p + 1)(p + 2)(p + 3),$$

and we have allowed for a finite contribution, with coefficient $\beta$, of the third counterterm, which is necessary to cancel a logarithmic divergence of the tensor perturbation. One gets the quantum effective action of the Yang–Mills theory on a general, perturbed FLRW geometry. For completeness, we also give the Einstein–Hilbert action of the perturbed four sphere,

$$S_{EH} = -\frac{3\Omega_4 r^2}{4\pi G} - \frac{3}{4\pi G} \int d^4x \sqrt{\hat{\gamma}} \psi$$

$$+ \frac{1}{16\pi G r^2} \int d^4x \sqrt{\hat{\gamma}} \left( \frac{3}{2} \psi \hat{\nabla}^2 \psi + 2\theta^{\mu\nu} \theta_{\mu\nu} - \frac{1}{4} \theta^{\mu\nu} \hat{\nabla}^2 \theta_{\mu\nu} \right).$$

We shall use these results in section III, where we discuss the instability of anomaly–induced inflation. But first, we return to the background evolution. In the next paragraph, we discuss a class of $O(4)$ invariant ‘real
tunnelling’ instanton solutions of the Starobinsky model \(4.79\) and study their role in the no boundary path integral for the wave function of an inflationary universe.

**Real Tunnelling Geometries**

It is easily seen that the total action is stationary under all perturbations \(h_{\mu\nu}\), if the background is a round four sphere with radius

\[
r_s^2 = \frac{N^2 G}{4\pi}.
\]

(4.85)

By slicing the four sphere at the equator \(\chi = \pi/2\) and writing \(\chi = \pi/2 - it\), it analytically continues into the Lorentzian to the de Sitter solution mentioned above, with the cosmological constant provided by the trace anomaly of the large \(N\) Yang–Mills theory.

Other compact, real instanton solutions of the form

\[
ds^2 = d\tau^2 + b^2(\tau)d\Omega_3^2
\]

(4.86)

were found in [Hawking et al. (2001)], by numerically integrating the Einstein equations, which can be obtained directly from the trace anomaly by using energy–momentum conservation. Imposing regularity at the North Pole (at \(\tau = 0\)) of the instanton leaves only the third derivative of the scale factor at the North Pole as an adjustable parameter. It is convenient to define dimensionless variables \(\tilde{\tau} = \tau/r_s\) and \(f(\tilde{\tau}) = b(\tau)/r_s\). For \(\alpha < 0\), there exists a second regular, compact ‘double bubble’ instanton, with \(f'''(0) = -2.05\), together with a one-parameter family of instantons with an irregular South Pole. For \(f'''(0) < -1\), the scale factor of the latter has two peaks. For \(-1 < f'''(0) < 0\) on the other hand, they are similar to the singular Hawking–Turok instantons that have been considered in the context of scalar field inflaton [Hawking and Turok (1998)].

The Lorentzian part of the real tunnelling saddle–points is obtained by analytically continuing the instanton metric across a hypersurface of vanishing extrinsic curvature. The double bubble instanton can be continued across its ‘equator’ to give a closed FLRW universe, or into an open universe by a double continuation across the South Pole. Our numerical studies show that the closed universe rapidly collapses and that the open space–time hyper-inflates, with the scale factor blowing up at a finite time. Similarly, the singular instantons can be continued into an open FLRW universe across \(\tau = 0\), by setting \(\tau = it\) and \(\Omega_3 = i\phi\). For \(f'''(0) < -1\)
this again gives hyper-inflation, but for $-1 < f''(0) < 0$ one gets a realistic inflationary universe. The four sphere solution as well as the singular instantons that are small perturbations of $S^4$ at the regular pole, are most interesting for cosmology, since they yield long periods of inflation.

Using the expression (4.83) for $W[\sigma(\chi)]$ and the relations

$$
\chi(\tau) = 2 \lim_{\epsilon \to 0} \tan^{-1} \left[ \tan(\epsilon/2) \exp \left( \int_{\epsilon}^{\tau} \frac{d\tau'}{b(\tau')} \right) \right], \quad B(\tau) = \frac{b(\tau)}{\sin(\chi)},
$$

one can numerically compute the action of the real tunnelling geometries [Hertog (2002)]. On an unperturbed FLRW background, conformal to the round four sphere, the total Euclidean action becomes [Hawking and Hertog (2002)]

$$
S^{(0)} = \frac{3N^2\Omega_3}{32\pi^2} \int d\chi \sin^3 \chi \left[ \frac{1}{3} (12(\log 2 + \sigma - \beta) - 3 + 6\sigma'^2 - \sigma'^4
- 4\sigma'^3 \cot \chi) - e^{2\sigma}(\sigma'^2 + 2) + 2\alpha(\sigma'' + 3\sigma' \cot \chi + \sigma'^2 - 2)^2 \right],
$$

where $\sigma = \log(B/r)$. On the round four sphere, $\sigma \to 0$, so the action reduces to

$$
S^{(0)} = \frac{3N^2\Omega_4}{32\pi^2} (8\alpha - 3 + 4(\log 2 - \beta)). \quad (4.87)
$$

We find that for all $\alpha < 0$ the regular double bubble instanton has much lower action than the four sphere. The singular double bubble instantons have divergent action, but the Hawking–Turok type instantons have finite action. For given $\alpha$, the action of the latter class depends on the third derivative of the scale factor at $\tau = 0$. This is the analogue of the situation in scalar field inflaton, where the action of the Hawking–Turok instantons depends on the value of the inflaton field at the North Pole. The action of the singular instantons tends smoothly to the $S^4$ action (4.87) as $f'''(0) \to -1$ and it decreases monotonically with increasing $f'''(0)$.

To summarize, we found a one–parameter family of finite–action, compact solutions of the Euclidean field equations that can be analytically continued across a space–like surface $\Sigma$ of vanishing curvature, to Lorentzian geometries that describe realistic inflationary universes. The condition on $\Sigma$ guarantees that a real solution of the Euclidean field equations is continued to a real Lorentzian space–time. The Euclidean region is essential, since there is no way to round off a Lorentzian geometry without introducing a boundary. What is the relevance then, in the context of the no boundary
proposal, of these real tunnelling geometries with regard to the problem of initial conditions in cosmology?

At least at the semiclassical level, the no boundary proposal gives a measure on the space of initial conditions for cosmology. The weight of each classical trajectory is approximately \( |\Psi|^2 \sim e^{-2S_R} \), where \( S_R \) is the real part of the Euclidean action of the solution. For real tunnelling solutions this comes entirely from the part of the manifold on which the geometry is Riemannian. The simplicity of this situation has led to the interpretation of the no boundary proposal as a bottom up theory of initial conditions. In particular, it has been argued that if a given theory allows different instantons, the no boundary proposal predicts our universe to be created through the lowest–action solution, since this would give the dominant contribution to the path integral. Applying this interpretation to trace anomaly driven inflation, one must conclude that the no boundary proposal predicts the creation of a hyper–inflating universe emerging from the double bubble instanton, or a nearly empty open universe that occurs by semiclassical tunnelling via a singular instanton with \( |f'''(0)| \) small.

The situation is similar in many theories of scalar field inflaton. Restricting attention to real tunnelling geometries, a bottom up interpretation of the no boundary proposal generally favors the creation of large space–times. One typically gets a probability distribution that is peaked around instantons in which the field at the surface of continuation is near the minimum of its potential, yielding very little inflation. Hence, the most probable universes are nearly empty open universes or collapsing closed universes, depending on the analytic continuation one considers. Weak anthropic arguments have been invoked to try to rescue the situation [Hawking and Turok (1998)], by weighing the \emph{a priori} no boundary probability with the probability of the formation of galaxies. However, for the most natural inflaton potentials, this still predicts a value of \( \Omega_0 \) that is far too low to be compatible with observations. Another attempt [Turok (2000)], based on introducing a volume factor that represents the projection onto the subset of states containing a particular observer, leads to eternal inflation at the Planck density, where the theory breaks down. In fact, invoking conditional probabilities is contrary to the whole idea of the no boundary proposal, which by itself specifies the quantum state of the universe.

Clearly the predictions of a bottom up interpretation of the no boundary proposal do not agree with observation. This is because it is an essentially classical interpretation, which is neither relevant nor correct for cosmology. The quantum origin of the universe implies its quantum state is given by
a path integral. Therefore, one must adopt a quantum approach to the problem of initial conditions, in which one considers the no boundary path integral (4.98) for a given quantum state of the universe. We shall apply such a quantum approach in section IV, to describe the origin of an inflationary universe, in theories like trace anomaly inflation. It turns out that the relevant saddle–points are not exactly real tunnelling geometries. Instead, one must consider complex saddle points, in which the geometry becomes gradually Lorentzian at late times [Hawking and Hertog (2002)].

4.3.6.2 Instability of Anomaly–Induced Inflation

Metric Perturbations

Two-point functions of metric perturbations can be computed directly from the no boundary path integral. One perturbatively evaluates the path integral around an $O(4)$ invariant instanton background to get the real–space Euclidean correlator, which is then analytically continued into the Lorentzian universe, where it describes the quantum fluctuations of the graviton field in the primordial de Sitter phase [Gratton and Turok (1999)] [Hertog and Turok (2000)]. The quantum state of the Lorentzian fluctuations is uniquely determined by the condition of regularity on the instanton [Hawking et al. (2001)]. Both scalar and tensor perturbations are given by a path integral of the form [Hawking and Hertog (2002)]

$$\langle h_{\mu\nu}(x)h_{\mu'\nu'}(x') \rangle \sim \int \mathcal{D}[h] \exp(-S^{(2)}h_{\mu\nu}(x)h_{\mu'\nu'}(x')), \quad (4.88)$$

where $S^{(2)}$ denotes the second order perturbation of the action

$$S = S_{EH} + S_{GH} + S_{R^2} + \tilde{W}, \quad (4.89)$$

with $\tilde{W}$ given by (4.82). For the scalars, eliminating the remaining gauge freedom introduces Faddeev–Popov ghosts. These ghosts supply a determinant $(\nabla^2 + 4)^{-1}$, which cancels a similar factor in the scalar action, rendering it second order. The action for the tensors $\theta_{\mu\nu}$ on the other hand is non-local and fourth order. Nevertheless, the metric perturbation and its first derivative should not be regarded as two independent variables, since this would lead to meaningless probability distributions in the Lorentzian [Hawking and Hertog (2001)]. Instead the path integral should be taken

$^{24}$The gauge freedom also leads to closed loops of Faddeev–Popov ghosts but they can be neglected in the large $N$ approximation.
over the fields $\theta_{\mu\nu}$ only \cite{25} to compute correlators of the form \eqref{4.88}. The Euclidean action for $\theta_{\mu\nu}$ is positive definite, so the path integral over all $\theta_{\mu\nu}$ converges and determines a well-defined Euclidean quantum field theory. One might worry that the higher derivatives would lead to instabilities in the Lorentzian. This is not the case, however, since the no boundary prescription to compute Lorentzian propagators by Wick rotation from the Euclidean, implicitly imposes the final boundary condition that the fields remain bounded, which eliminates the runaways \cite{Hawking et al. (2001), Hawking and Hertog (2001)}.

The path integral \eqref{4.88} is Gaussian, so the correlation functions can be read off from the perturbed action and equation \eqref{4.85}:

$$\langle \psi(x)\psi(x') \rangle = \frac{32\pi^2 r_s^4}{3|\alpha|N^2} \left(-\nabla^2 + 1/2\alpha\right)^{-1}, \quad \text{and} \quad (4.90)$$

$$\langle \theta_{\mu\nu}(x)\theta_{\mu'\nu'}(x') \rangle = \frac{128\pi^2 r_s^4}{N^2} \sum_{p=2}^{\infty} \frac{W^{(p)}_{\mu\nu\mu'\nu'}(x,x')}{p^2 + 3p + 6 + \Psi(p) - 4\alpha p + 3}, \quad (4.91)$$

where the bitensor $W^{(p)}_{\mu\nu\mu'\nu'}(x,x')$ is defined as the usual sum over degenerate rank–2 harmonics on the four sphere.

The scalar 2–point function \eqref{4.90} is just the propagator of a particle with physical mass $m^2 = (2\alpha r_s^2)^{-1}$. Since we are assuming $\alpha < 0$, we have $m^2 < 0$ so this particle is a tachyon, which is the perturbative manifestation of the Starobinsky instability. Making $\alpha$ more negative, makes the tachyon mass squared less negative, and therefore weakens the instability. Indeed, the number of efoldings in the primordial de Sitter phase emerging from the four sphere instanton is given by $N_{\text{efolds}} \sim 12|\alpha| (\log N - 1)$. Therefore, in the interesting regime, we have $-m^2 \ll m_{pl}^2$, so semiclassical gravity should be a good approximation.

This result sheds light on the problem of initial conditions in trace anomaly inflation. One can think of the non derivative term in the scalar correlator as a potential $V(\psi)$, with the unperturbed de Sitter solution at $\psi = 0$ at the maximum. If $|\alpha|$ is not too small, then the top of the potential is sufficiently flat, so that the lowest–action regular instanton is a homogeneous Hawking–Moss instanton \cite{Hawking and Moss (1982)}.

\footnote{This means one loses unitarity. However, probabilities for observations tend towards those of the second order theory, as the coefficients of the fourth order terms in the action tend to zero. Hence unitarity is restored at the low energies that now occur in the universe.}
with $\psi$ constant at the top. Since the instability of the de Sitter phase is characterized entirely by the coefficient $\alpha$ of the $R^2$ counterterm, this means the problem of initial conditions in anomaly–induced inflation is similar to the corresponding problem in many theories of scalar field inflaton, where one ought to explain why the inflation starts initially at the top of the hill.

**Homogeneous Fluctuations**

The most interesting instantons in both trace anomaly driven inflation as well as most theories of scalar field inflaton possess a homogeneous fluctuation mode which decreases their action [Hawking et al. (2001); Gratton and Turok (2001)]. The presence of such a negative mode is the perturbative manifestation of the conformal factor problem. Indeed, since the conformal factor problem is closely related to the instability of gravity under gravitational collapse, one expects instantons that are appealing from a cosmological perspective, to possess a negative mode.

Writing the scalar propagator (4.90) on the four sphere instanton in momentum space gives [Hawking and Hertog (2002)]

$$
\langle \psi(x)\psi(x') \rangle = \frac{32\pi^2 r_s^4}{3|\alpha|N^2} \sum_{p=0}^{\infty} \frac{W^{(p)}(\mu(x,x'))}{p(p+3) + m^2},
$$

where the biscalar $W^{(p)}$ equals the usual sum over degenerate scalar harmonics on the four sphere with eigenvalue $\lambda_p = -p(p+3)$ of the Laplacian.

There are many negative modes if $-1/8 < \alpha < 0$. This is usually the perturbative indication of the existence of a lower-action instanton solution. For instance, in scalar field inflaton with a double well potential, the Hawking–Moss instanton possesses several negative modes if $V_{\phi\phi}/H^2 < -4$, which is precisely the condition for the existence of a lower–action Coleman–De Luccia instanton that straddles the maximum. On the other hand, if $\alpha < -1/8$ in (4.92) then only the homogeneous ($p = 0$) negative mode remains, which is again similar to the well–known negative homogeneous mode of the Hawking–Moss instanton in theories with a scalar potential that is sufficiently flat.

The presence of a physical negative mode supports the interpretation of an instanton as describing the decay of an unstable state through semiclassical tunnelling [Coleman (1980)]. On the other hand, it has been argued that it questions its use in the no boundary path integral to define the ini-
tial quantum state of the universe\cite{Gratton:2001}. Within the semiclassical approximation, however, it is more appropriate to project out the negative mode, since the semiclassical approach is based on the assumption that the path integral can be expanded around solutions of the classical field equations.

The conclusions of\cite{Gratton:2001} are based on a perturbation calculation around compact, real instanton backgrounds, that does not take into account the wave function of interest. One expects, however, the configuration specifying the quantum state of the Lorentzian universe to project out the negative mode from the perturbation spectrum. Consider for example the wave function of a universe described by a 3–sphere with radius $R^2 = V_0/3$ and field $\phi = 0$, in a theory of gravity coupled to a single scalar field with potential $V_0(1 - \phi^2)^2$. In the semiclassical approximation, this is given by half of a Hawking–Moss instanton with the field constant at the top of the potential. Obviously, this solution has no negative mode, since the boundary condition on the 3–sphere $\Sigma$ removes the lowest eigenvalue solution of the Schrödinger equation for the perturbations. Since the negative mode corresponds to a homogeneous fluctuation, this is probably true also for large 3–spheres in the Lorentzian regime. Therefore, one expects that in the top down approach to cosmology, where the quantum state of the universe is taken into account, the negative mode is automatically projected out.

Quantum Matter and the Microwave Background

Before discussing the top down approach in more detail, we pause to briefly comment on some of the characteristic predictions for observations of trace anomaly inflation.

To extract accurate predictions for the cosmic microwave anisotropies, one must evolve the perturbations through the Starobinsky instability, to get initial conditions for the inhomogeneities during the radiation and matter eras. Details of this calculation will be presented elsewhere\cite{Hawking:2002}, but some interesting features of the microwave temperature anisotropies predicted by anomaly–induced inflation, can be ex-

\footnote{In scalar field inflaton, one can view the singular Hawking–Turok instantons as constrained instantons, with additional data specified on an internal boundary. For some theories, the constraint introduced in\cite{Kirklin:2000} to resolve the singularity, also removes the negative mode, at least perturbatively\cite{Gratton:2001}. However, it does not remove the instability non–perturbatively and for the most obvious potentials, the lowest–action constrained instanton gives very little inflation.}
trated from the correlators (4.90) and (4.91) in the primordial de Sitter era. As can be seen from (4.91), the quantum matter couples to the tensors. Starobinsky [Starobinsky (1983)] and Vilenkin [Vilenkin (1985)] assumed that the amplitude of primordial gravity waves was not significantly altered by the quantum matter loops. This assumption can now be examined using AdS/CFT, which has allowed us to include the effect of the Weyl $^2$ counterterm and the non-local part of the matter effective action. We find that at small scales, matter fields dominate the tensor propagator and make it decay like $p^4 \log p$. In other words, the CFT appears to give space-time a rigidity on small scales, an example of how quantum loops of matter can change gravity at short distances. In fact, this suppression should occur even if inflation is not driven by the trace anomaly, since we observe a large number of matter fields, whose effective action is expected to dominate the propagator on small scales.

Secondly, both the higher derivative counterterms and the matter fields introduce anisotropic stress, which is an important difference with scalar field inflaton. This can be seen from decomposing the tensors $\theta_{\mu\nu}$ into a scalar $\phi$ and tensor $t_{ij}$ under $O(4)$. The former is the difference between the two potentials in the Newtonian gauge and corresponds to anisotropic stress. Typically reheating at the end of anomaly-induced inflation leads to creation of particles that are not in thermal equilibrium with the photon-baryon fluid, so one expects some anisotropic stress to survive during the radiation era. To make more precise predictions, however, a better understanding is required of the (probably time-dependent) values of the coefficients $\alpha$ and $\beta$ of the higher derivative counterterms in the theory.

Finally, we should mention that for the tensor propagator the higher derivative terms also give rise to poles in the complex $p-$plane. These are harmless, however, since the contour obtained from the Euclidean goes around the complex poles [Hawking et al. (2001)]. In other words, defining the theory in the Euclidean, implicitly removes the instabilities associated with the complex poles, like a final boundary condition removes the runaway solution of the classical radiation reaction force [Hawking and Hertog (2001)].

4.3.6.3 Origin of Inflation

We have seen that the predictions of the bottom up approach to the problem of initial conditions in inflation do not agree with observation. This is because it is based on an essentially classical picture, in which one assumes
some initial condition for the universe and evolves it forward in time. The
quantum origin of our universe, however, means that its wave function is
determined by a path integral, in which one sums over all possible histories
that lead to a given quantum state, together with some suitable boundary
conditions on the paths. This naturally leads to a top down view of the
universe. In a top down context, rather than comparing the relative proba-
bilities of different semiclassical geometries, one looks for the most probable
evolution that leads to a certain outcome.

We now apply the quantum top down interpretation of the no boundary
proposal to study the origin of an inflationary universe, in theories where
the instability of the inflationary phase can be described in terms of a single
scalar field with an effective potential that has a local maximum. As shown
above, this includes trace anomaly driven inflation, since the emergence of
an anomaly driven inflationary universe is very similar to the creation of
an exponentially expanding universe in theories of new inflation.

We consider a model consisting of gravity coupled to a single scalar
field, with a double well potential

\[ V(\phi) = A(1 - \frac{C}{2} \phi^2)^2 \] (with \( A, C > 0 \)).

For \( C < 2/3 \), the potential has a maximum at \( \phi = 0 \) with \( V_{,\phi\phi}/V \) suf-
ficiently low so that there exists no Coleman–De Luccia instanton, but
only a Hawking–Moss instanton with \( \phi = 0 \) everywhere on top of the hill.
Implementing a top down approach, we consider the quantum amplitudes
\( \Phi[\tilde{h}, K, \phi_\Sigma] \) for different conformal 3-geometries \( \tilde{h} \) with trace
\( K \) of the second fundamental form, on an expanding surface \( \Sigma \) during inflation. Ac-
cording to the no boundary proposal, the defining path integral should be
taken over all compact Riemannian geometries that induce the prescribed
configuration on \( \Sigma \).

In the \( K \)-representation, the Euclidean action is given by [Hawking
and Hertog (2002)]

\[ S = -\frac{1}{2\kappa} \int d^4x \sqrt{g} R - \frac{1}{3\kappa} \int d^3x \sqrt{h} K + \int d^4x \sqrt{g} \left( \frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi + V(\phi) \right). \] (4.93)

The usual wave function \( \Psi[\mathbf{h}, \phi_\Sigma] \) is obtained from \( \Phi[\tilde{h}, K, \phi_\Sigma] \) by an inverse
Laplace transform,

\[ \Psi[\mathbf{h}, \phi_\Sigma] = \int_{\Gamma} D \left[ \frac{K}{4i\kappa} \right] \exp \left[ \frac{2}{3\kappa} \int d^3x \sqrt{h} K \right] \Phi[\tilde{h}, K, \phi_\Sigma] \]

where the contour \( \Gamma \) runs from \(-i\infty\) to \(+i\infty\).

Within the semiclassical approximation, the no boundary wave function
is approximately given by the saddle–point contributions. Restricting
attention to saddle–points that are invariant under the action of an $O(4)$ isometry group, the instanton metric can be written as
\begin{equation}
    ds^2 = d\tau^2 + b^2(\tau)d\Omega^2_3,
\end{equation}
and the Euclidean field equations read
\begin{align}
    \phi'' &= -K\phi' + V_{,\phi}, \\
    K' + K^2 &= -(\phi^2_{,\tau} + V),
\end{align}
where $\phi' = \phi_{,\tau}$ and $K = 3b_{,\tau}/b$. The Lorentzian trace $K_L = -3a/a$ is obtained by analytic continuation. We first calculate the wave function for real $K$, and then analytically continue to imaginary, or Lorentzian $K_L = -iK$.

At the semiclassical level, there are two contributions to the given amplitude. For small $\phi_\Sigma$ and any Euclidean $K$, there always exists a non–singular, Euclidean $O(4)$ invariant solution of the field equations, with the prescribed boundary conditions. This solution is part of a deformed sphere, or Hawking–Turok instanton. In the approximation $K = 3H \cot(\theta_H)$, with $H^2 = A/3$, and $V(\phi) \sim A(1 - C\phi^2)$, the solution of (4.106) is given by
\begin{equation}
    \phi = \phi_\Sigma \frac{2F_1(3/2 + q, 3/2 - q, 2, z(K))}{2F_1(3/2 + q, 3/2 - q, 2, z(K_\Sigma))},
\end{equation}
where
\begin{align}
    q &= \sqrt{9/4 + 6C}, \\
    z(K) &= \frac{1}{2} \left[ 1 - \frac{K}{(A^2 + K^2)^{1/2}} \right].
\end{align}

At the South Pole $K \to +\infty$, so in the instanton the scalar field slowly rolls up the hill from its value at the regular South Pole to the prescribed value $\phi_\Sigma$ on the 3-sphere with trace $K_\Sigma$. The weight of the Hawking–Turok geometry in the no boundary path integral for the wave function $\Phi[K, \phi_\Sigma]$ is approximately given by [Hawking and Hertog (2002)]
\begin{align}
    S[K_\Sigma, \phi_\Sigma] &= -\frac{1}{3\kappa} \int d^3x \sqrt{\eta}K - \int d^4x \sqrt{\eta}V(\phi) \\
    &= -\frac{12\pi^2}{A} \left[ 1 - \frac{K}{(A^2 + K^2)^{1/2}} \right] - \frac{24\pi^2C}{A^2(1 - C)} \phi_\Sigma^2 z^2(K_\Sigma) \times \\
    &\quad \left[ 1 - 2z(K_\Sigma) + 3C(1 - z(K_\Sigma)) \right] \left( z(K_\Sigma) + \frac{2 - 3C}{1 + 3C} \right) F' \left[ z(K_\Sigma) \right].
\end{align}

For small $\phi_\Sigma$, there is a second semiclassical contribution to the wave function, coming from universes that are created via an $O(5)$ symmetric Hawking–Moss instanton with $\phi$ constant at the top of the hill, but in
which a quantum fluctuation disturbs the field, causing it to run down to its prescribed value $\phi_\Sigma$ at the 3–sphere boundary with trace $K_\Sigma$. It follows that for $K_\Sigma = 0$, the action of the Hawking–Turok geometry is more negative than the action of the Hawking–Moss instanton. This would seem to suggest that the universe is least likely to start at the top of the hill. However, we are not interested in the amplitude for a Euclidean space–time, but in the no boundary wave function of a Lorentzian expanding universe.

Within the regime in which $\phi$ remains small over the whole geometry, one can derive the amplitude in the Lorentzian from the result for $S[K_\Sigma, \phi_\Sigma]$, by analytic continuation into the complex $K$–plane. In a Lorentzian universe, Euclidean $K$ is pure imaginary, $K_L = -iK$. Since the action is invariant under diffeomorphically related contours in the complex $\tau$–plane, we may deform the contour into one with straight sections, along the real and imaginary $K$–axis. The real part of the action for the Hawking–Moss instanton is constant on the imaginary $K$–axis, unlike the action for the Hawking–Turok geometry. According to the no boundary proposal, the relative probability of both geometries is given by

$$P[K_L, \phi_\Sigma] = \frac{A_{HM}^2}{A_{HT}^2} e^{-2\Re[\Delta S]},$$

where $\Delta S = S_{HM} - S_{HT}$. The prefactors account for small fluctuations around the classical solutions and can be neglected for small $\phi$.

As mentioned above, this scenario is realised in trace anomaly driven inflation. The unperturbed de Sitter solution (4.85) in anomaly–induced inflation emerges from the Hawking–Moss geometry, while the inhomogeneous Hawking–Turok evolution corresponds to one of the singular instantons discussed above. The field configuration on $\Sigma$ determines the third derivative of the scale factor at the regular South Pole, or equivalently the initial value of the order parameter $\phi$ governing the instability. For $\alpha < -1/8$, the instability of the de Sitter phase is sufficiently weak, so that the universe is most likely to start at the top, in an unstable de Sitter state. This result also justifies presented calculation of metric perturbations, which were based on a perturbative expansion of the path integral about the round four sphere.

Finally, we should mention that because we are interested in real matter fields on $\Sigma$, the analytic continuation into the complex $K$–plane means $\phi$ must be complex in the bulk of the instanton. More precisely, at the South Pole, we must have

$$\Im[\phi] = \phi_\Sigma \Im[F(z_\Sigma)] / \Re[F(z_\Sigma)].$$
This has no physical meaning though, since the stationary phase approximation is just a mathematical construction to evaluate the path integral over real $\phi$.

For more details on top–down inflation, see [Hawking and Hertog (2002)].

### 4.3.7 Cosmology in the String Landscape

In this final section on cosmological dynamics, we present the seminal paper by S. Hawking and T. Hertog [Hawking and Hertog (2006)], proposing a top–down/no–boundary approach to quantum cosmology in the string landscape. They put forward a framework for cosmology that combines the string landscape with no boundary initial conditions. In this framework, amplitudes for alternative histories for the universe are calculated with final boundary conditions only. This leads to a top down approach to cosmology, in which the histories of the universe depend on the precise question asked.

The authors study the observational consequences of no boundary initial conditions on the landscape, and outline a scheme to test the theory. This is illustrated in a simple model landscape that admits several alternative inflationary histories for the universe. Only a few of the possible vacua in the landscape will be populated. They also discuss in what respect the top down approach differs from other approaches to cosmology in the string landscape, like eternal inflation.

Now, it seems likely that string theory contains a vast ensemble of stable and meta–stable vacua, including some with a small positive effective cosmological constant [Bousso and Polchinski (2000)] and the low energy effective field theory of the *Standard Model*. Recent progress on the construction of meta–stable de Sitter vacua [Kachru et al. (2003)] lends further support to the notion of a *string landscape* [Susskind (2003)], and a statistical analysis gives an idea of the distribution of some properties among the vacua [Douglas (2003)]. But it has remained unclear what is the correct framework for cosmology in the string landscape. There are good reasons to believe, however, that a proper understanding of the cosmological dynamics will be essential for the landscape to be predictive [Banks et al. (2004)].

Recall that in particle physics, one usually computes $S$–matrix elements. This is useful to predict the outcome of laboratory experiments, where one prepares the initial state and measures the final state. It could be viewed as a bottom–up approach to physics, in which one evolves forward in time a particular initial state of the system. The predictability of this approach
arises from and relies upon the fact that one has control over the initial state, and that experiments can be repeated many times to gain statistically significant results.

But cosmology poses questions of a very different character. In our past there is an epoch of the early universe when quantum gravity was important. The remnants of this early phase are all around us. The central problem in cosmology is to understand why these remnants are what they are, and how the distinctive features of our universe emerged from the Big–Bang. Clearly it is not an S–matrix that is the relevant observable\(^{27}\) for these predictions, since we live in the middle of this particular experiment. Furthermore, we have no control over the initial state of the universe, and there is certainly no opportunity for observing multiple copies of the universe.

In fact if one does adopt a bottom–up approach to cosmology, one is immediately led to an essentially classical framework, in which one loses all ability to the central question of cosmology: why our universe is the way it is. In particular a bottom–up approach to cosmology either requires one to postulate an initial state of the universe that is carefully fine–tuned [Gasperini and Veneziano (1993)], as if prescribed by an outside agency, or it requires one to invoke the notion of eternal inflation [Vilenkin (1983)], which prevents one from predicting what a typical observer would see.

Here we put forward a different approach to cosmology in the string landscape, based not on the classical idea of a single history for the universe but on the quantum sum over histories [Hartle (1995a)]. We argue that the quantum origin of the universe naturally leads to a framework for cosmology where amplitudes for alternative histories of the universe are computed with boundary conditions at late times only. We thus envision a set of alternative universes in the landscape, with amplitudes given by the no–boundary path integral [Hartle and Hawking (1983)].

The measure on the landscape provided by no–boundary initial conditions allows one to derive predictions for observations. This is done by evaluating probabilities for alternative histories that obey a set of constraints at late times. The constraints provide information that is supplementary to the fundamental laws and act as a selection principle. In particular, they select the subclass of histories that contribute to the amplitude of interest. One then identifies alternatives within this subclass that have probabilities

---

\(^{27}\)See [Banks and Fischler (2001); Witten (2001); Bousso (2005); Giddings et al. (2005)] for recent work on the existence and the construction of observables in cosmological space–times.
near one. These include, in particular, predictions of future observations. The framework proposed here is thus more like a top–down approach to cosmology, where the histories of the universe depend on the precise question asked [Hawking and Hertog (2006)].

4.3.7.1 Quantum State

In cosmology one is generally not concerned with observables at infinity or with properties of the entire 4–geometry, but with alternatives in some finite region in the interior of the space–time. The amplitudes for these more restricted sets of observables are obtained from the amplitudes of four dimensional metric and matter field configurations, by integrating over the unobserved quantities. A particularly important case is the amplitude of finding a compact space–like surface $S$ with induced 3–metric $g_{ij}$ and matter field configuration $\phi$ [Hawking and Hertog (2006)],

$$\Psi[g^3, \phi] \sim \int_C [\mathcal{D}g][\mathcal{D}\phi] e^{iS[g, \phi]}.$$ (4.97)

Here the path integral is taken over the class $C$ of space–times which agree with $g_{ij}$ and $\phi$ on a compact boundary $S$. The quantum state of the universe is determined by the remaining specification of the class C.

Usually one sums over histories that have an initial and a final boundary. This is useful for the computation of $S$–matrix elements to predict the outcome of laboratory experiments, where one prepares the initial state and measures the final state. It is far from clear, however, that this is the appropriate setup for cosmology, where one has no control over the initial state, and no opportunity for observing multiple copies of the universe. In fact, if one does apply this approach to cosmology one is naturally led to an essentially classical picture, in which one simply assumes the universe began and evolved in a way that is well defined and unique.

The so–called pre–Big–Bang cosmologies [Gasperini and Veneziano (1993)] are examples of models that are based on a bottom–up approach. In these models one specifies an initial state on a surface in the infinite past and evolves this forward in time. A natural choice for the initial state would be flat space, but that would obviously remain flat space. Thus one instead starts with an unstable state in the infinite past, tuned carefully in order for the Big–Crunch/Big–Bang transition to be smooth and the path integral to be peaked around a single semi–classical history. Several explicit solutions of such bouncing cosmologies
have been found in various mini–superspace approximations [Karczmarek and Strominger (2004)]. It has been shown, however, using several different techniques, that solutions of this kind are unstable [Liu et al. (2002); Hertog and Horowitz (2004)]. In particular, one finds that generic small perturbations at early times (or merely taking in account the remaining degrees of freedom) dramatically change the evolution near the transition. Rather than evolving towards an expanding semi–classical universe at late times, one generically produces a strong curvature singularity. Hence the evolution of pre–Big–Bang cosmologies always includes a genuinely quantum gravitational phase, unless the initial state is extremely fine–tuned. It is therefore more appropriate to describe these cosmologies by a path integral in quantum cosmology, and not in terms of a single semi–classical trajectory. The universe will not have a single history but every possible history, each with its own probability.

In fact, the quantum state of the universe at late times is likely to be independent of the state on the initial surface. This is because there are geometries in which the initial surface is in one universe and the final surface in a separate disconnected universe. Such metrics exist in the Euclidean regime, and correspond to the quantum annihilation of one universe and the quantum creation of another. Moreover, because there are so many different possible universes, these geometries dominate the path integral. Therefore even if the path integral had an initial boundary in the infinite past, the state on a surface $S$ at late times would be independent of the state on the initial surface. It would be given by a path integral over all metric and matter field configurations whose only boundary is the final surface $S$. But this is precisely the no–boundary quantum state [Hartle and Hawking (1983)]

$$\Psi[g^3, \phi] \sim \int_C [\mathcal{D}g][\mathcal{D}\phi] e^{-S_E[g, \phi]},$$

(4.98)

where the integral is taken over all regular geometries bounded only by the compact 3–geometry $S$ with induced metric $g^3_{ij}$ and matter field configuration $\phi$. The Euclidean action $S_E$ is given by\(^{28}\)

$$S_E = -\frac{1}{2} \int d^4x \sqrt{g} (R + L(g, \phi)) - \int_S d^3x \sqrt{g^3} K,$$

(4.99)

where $L(g, \phi)$ is the matter Lagrangian.

\(^{28}\)We have set $8\pi G = 1$. 

Quantum Leap
One expects that the dominant contributions to the path integral will come from saddle points in the action. These correspond to solutions of the Einstein equations with the prescribed final boundary condition. If their curvature is bounded away from the Planck value, the saddle point metric will be in the semi-classical regime and can be regarded as the most probable history of the universe. Saddle point geometries of particular interest include geometries where a Lorentzian metric is rounded off smoothly in the past on a compact Euclidean instanton. Well known examples of such geometries are the Hawking–Moss (HM) instanton [Hawking and Moss (1982)] which matches to Lorentzian de Sitter space, and the Coleman–De Luccia (CdL) instanton [Coleman (1980)], which continues to an open FLRW universe. The former occurs generically in models of gravity coupled to scalar fields, while the latter requires a rather fine-tuned potential.

The usual interpretation of these geometries is that they describe the decay of a false vacuum in de Sitter space. However, they have a different interpretation in the no–boundary proposal [Gratton and Turok (1999)]. Here they describe the beginning of a new, independent universe with a completely self-contained ‘no–boundary’ description\textsuperscript{29}. By this we mean, in particular, that the expectation values of observables that are relevant to local observers within the universe can be unambiguously computed from the no–boundary path integral, without the need for assumptions regarding the pre–bubble era. The original de Sitter universe may continue to exist, but it is irrelevant for observers inside the new universe. The no–boundary proposal indicates, therefore, that the pre-bubble inflating universe is a redundant theoretical construction.

It is appealing that the no–boundary quantum state (4.98) is computed directly from the action governing the dynamical laws. There is thus essentially a single theory of dynamics and of the quantum state. It should be emphasized however that this remains a proposal for the wave function of the universe. We have argued it is a natural choice, but the ultimate test is whether its predictions agree with observations.

4.3.7.2 Prediction in Quantum Cosmology

Quantum cosmology aims to identify which features of the observed universe follow directly from the fundamental laws, and which features can be

\textsuperscript{29}The interpretation of these saddle point geometries is in line with their interpretation that follows from holographic reasoning, as described e.g., in [Dyson et al. (2002a)]. Some of our conclusions, however, differ from [Dyson et al. (2002a)].
understood as consequences of quantum accidents or late time selection effects. In no–boundary cosmology, where one specifies boundary conditions at late times only, this program is carried out by evaluating probabilities for alternative histories that obey certain constraints at the present time. The final boundary conditions provide information that is supplementary to the fundamental laws, which selects a subclass of histories and enables one to identify alternatives that (within this subclass) have probabilities near one. In general the probability for an alternative \( \alpha \), given \( H, \Psi \) and a set of constraints \( \beta \), is given by \(^{[\text{Hawking and Hertog (2006)}]}\)

\[
p(\alpha|\beta, H, \Psi) = \frac{p(\alpha, \beta|H, \Psi)}{p(\beta|H, \Psi)}. \tag{4.100}
\]

The conditions \( \beta \) in (4.100) generally contain environmental selection effects, but they also include features that follow from quantum accidents in the early universe \(^{[30]}\).

A typical example of a condition \( \beta \) is the dimension \( D \) of space. For good reasons, one usually considers string compactifications down to three space dimensions. However, there appears to be no dynamical reason for the universe to have precisely four large dimensions. Instead, the no–boundary proposal provides a framework to calculate the quantum amplitude for every number of spatial dimensions consistent with string theory. The probability distribution of various dimensions for the universe is of little significance, however, because we have already measured we live in four dimensions. Our observation only gives us a single number, so we cannot tell from this whether the universe was likely to be four dimensional, or whether it was just a lucky chance. Hence as long as the no–boundary amplitude for three large spatial dimensions is not exactly zero, the observation that \( D = 3 \) does not help to prove or disprove the theory. Instead of asking for the probabilities of various dimensions for the universe, therefore, we might as well use our observation as a final boundary condition and consider only amplitudes for surfaces \( S \) with three large dimensions. The number of dimensions is thus best used as a constraint to restrict the class of histories that contribute to the path integral for a universe like ours. This restriction allows one to identify definite predictions for future observations.

The situation with the low energy effective theory of particle interactions may well be similar. In string theory this is the effective field theory for the modular parameters that describe the internal space. It is well known that

\(^{[30]}\)These are quantum accidents that became ‘frozen’, leaving an imprint on the universe at late times.
string theory has solutions with many different compact manifolds. The corresponding effective field theories are determined by the topology and the geometry of the internal space, as well as the set of fluxes that wind the 3–cycles. Furthermore, for each effective field theory the potential for the moduli typically has a large number of local minima. Each local minimum of the potential is presumably a valid vacuum of the theory. These form a landscape [Susskind (2003)] of possible stable or meta–stable states for the universe at the present time, each with a different theory of low energy particle physics.

In the bottom–up picture it is thought that the universe begins with a grand unified symmetry, such as $E_8 \times E_8$. As the universe expands and cools the symmetry breaks to the Standard Model, perhaps through intermediate stages. The idea is that string theory predicts the pattern of breaking, and the masses, couplings and mixing angles of the Standard Model. However, as with the dimension of space, there seems to be no particular reason why the universe should evolve precisely to the internal space that gives the Standard Model. It is therefore more useful to compute no–boundary amplitudes for a space–like surface $S$ with a given internal space. This is the top–down approach, where one sums only over the subclass of histories which end up on $S$ with the internal space for the Standard Model [Hawking and Hertog (2006)].

We now turn to the predictions $\alpha$ we can expect to derive from amplitudes like (4.100). We have seen that the relative amplitudes for radically different geometries are often irrelevant. By contrast, the probabilities for neighboring geometries are important. The most powerful predictions are obtained from the relative amplitudes of nearby geometries, conditioned on various discrete features of the universe. This is because these amplitudes are not determined by the selection effects of the final boundary conditions. Rather, they depend on the quantum state $|\Psi\rangle$ itself.

Neighboring geometries correspond to small quantum fluctuations of continuous quantities, like the temperature of the cosmic microwave background (CMB) radiation or the expectation values of the string theory moduli in a given vacuum. In inflationary universes these fluctuations are

---

$^{31}$An extension of the bottom–up approach invokes the notion of eternal inflation to accommodate the possibility that the position in the moduli space falls into different minima in different places in space, leading to a mosaic structure for the universe. The problem with this approach is that one cannot predict what a typical local observer within such a universe would see.
amplified and stretched, generating a pattern of spatial variations on cosmological scales in those directions of moduli space that are relatively flat. The spectra depend on the quantum state of the universe. Correlators of fluctuations in the no–boundary state can be calculated by perturbatively evaluating the path integral around instanton saddle points [Gratton and Turok (1999)]. In general if $P(x_1)$ and $Q(x_2)$ are two observables at $x_1$ and $x_2$ on a final surface $S$, then their correlator is formally given by the following integral over a complete set of observables $O(x)$ on $S$ [Gratton and Turok (1999)],

$$\langle P(x_1)Q(x_2) \rangle \sim \sum_B \int [D O(S)] \Psi_B [O] \Psi_B [O]^* P(x_1)Q(x_2). \quad (4.101)$$

Here the sum is taken over backgrounds $B$ that satisfy the prescribed conditions on $S$. The amplitude $\Psi_B$ for fluctuations about a particular background geometry $(\bar{g}, \bar{\phi})$ is given by

$$\Psi_B [g, \phi] \sim e^{-S_0(\bar{g}, \bar{\phi})} \int [D \delta g][D \delta \phi] e^{-S_2[\delta g, \delta \phi]}, \quad (4.102)$$

where the metric $g = \bar{g} + \delta g$ and the fields $\phi = \bar{\phi} + \delta \phi$. The $C_l$'s of the CMB temperature anisotropies are classic examples of observables that can be calculated from correlators like this. Whilst the full correlator generally involves a sum over several saddle points, for most practical purposes only the lowest action instanton matters.

In no–boundary backgrounds like the HM geometry, where a real Euclidean instanton is matched onto a real Lorentzian metric, one can find the correlators by first calculating the 2–point functions in the Euclidean region. The Euclidean correlators are then analytically continued into the Lorentzian region, where they describe the quantum mechanical vacuum fluctuations of the various fields in the state determined by no–boundary initial conditions. The path integral unambiguously specifies boundary conditions on the Euclidean fluctuation modes. This essentially determines a reflection amplitude $R(k)$, where $k$ is the wave–number, which depends on the instanton geometry. The spectra in the Lorentzian, and in particular the primordial gravitational wave spectrum [Gratton et al. (2000)], depend on the instanton background through $R(k)$.

\[32\] Spatial variations of coupling constants from scalar moduli field fluctuations generate large scale iso–curvature fluctuations in the matter and radiation components [Kofman (2005)].
The relative amplitudes of neighboring geometries can thus be used to predict, from first principles, the precise shape of the primordial fluctuation spectra that we observe. This provides a test of the no–boundary proposal and, more generally, an observational discriminant between different proposals for the state of the universe, because the spectra contain a signature of the initial conditions [Hawking and Hertog (2006)].

4.3.7.3 Anthropic Reasoning

Recall that in general anthropic reasoning [Barrow and Tipler (1986)] aims to explain certain features of our universe from our existence in it. One possible motivation for this line of reasoning is that the observed values and correlations of certain parameters in particle physics and cosmology appear necessary to ensure life emerges in our universe. If this is indeed the case it seems reasonable to suppose that certain environmental selection effects need to be taken in account in the calculation of probabilities for observations.

It has been pointed out many times, however, that anthropic reasoning is meaningless if it is not implemented in a theoretical framework that determines which parameters can vary and how they vary. Top down cosmology, by combining the string landscape with the no–boundary proposal, provides such a framework. The anthropic principle is implemented in the top–down approach by specifying a set of conditions $\beta$ in (4.100) that select the subclass of histories where life is likely to emerge. More specifically, anthropic reasoning in the context of top–down cosmology amounts to the evaluation of conditional probabilities like [Hawking and Hertog (2006)]

$$p(\alpha | O, H, \Psi), \quad (4.103)$$

where $O$ represents a set of conditions that are required for the appearance of complex life. The utility and predictability of anthropic reasoning depends on how sensitive the probabilities (4.103) are to the inclusion of $O$. Anthropic reasoning is useful and predictive only if (4.103) is sharply peaked around the observed value of $\alpha$, and if the a priori theoretical probability $p(\alpha | H, \Psi)$ itself is broadly distributed [Hartle (2005a)].

Anthropic reasoning, therefore, can be naturally incorporated in the top–down approach. In particular it may provide a qualitative understanding for the origin of certain conditions $\beta$ that one finds are useful in top–down cosmology. Consider the number of dimensions of space, for example. We have argued that this is best used as a final constraint, but the top–
down approach itself does not explain why this particular property of the universe cannot be predicted from first principles. In particular, the top-down argument does not depend on whether four dimensions is the only arena for life. Rather, it is that the probability distribution over dimensions is irrelevant, because we cannot use our observation that $D = 3$ to falsify the theory. But it may turn out that anthropically weighted probabilities (4.103) are always sharply peaked around $D = 3$. In this case one can essentially interpret the number of dimensions as an anthropic requirement, and it would be an example where anthropic reasoning is useful to understand why one needs to condition on the number of dimensions in top-down cosmology.

We emphasize, however, that the top-down approach developed here goes well beyond conventional anthropic reasoning. Firstly, the top-down approach gives \textit{a priori} probabilities that are more sharply peaked, because it adopts a concrete prescription for the quantum state of the universe – as opposed to the usual assumption that predictions are independent of $\Psi$. Hence the framework we propose is more predictive than conventional anthropic reasoning.\footnote{Anthropic selection effects have been used to constrain the value of the cosmological constant [Weinberg (1987)], and the dark matter density [Linde (1988)]. In these studies it is assumed, however, that the \textit{a priori} probability distributions are independent of the state of the universe. This reduces the predictability of the calculations, and could in fact be misleading.}

Top down cosmology is also more general than anthropic reasoning, because there is a wider range of selection effects that can be quantitatively taken in account. In particular the conditions $\beta$ that are supplied in (4.100) need not depend on whether they are necessary for life to emerge. The set of conditions generally includes environmental selection effects similar to anthropic requirements, but it also includes chance outcomes of quantum accidents in the early universe that became frozen. The latter need not be relevant to the emergence of life. Furthermore, they cannot be taken in account by simply adding an \textit{a posteriori} selection factor proportional to the number density of some reference object, because they change the entire history of the universe!

\subsection*{4.3.7.4 Models of Inflation}

How can one get a nonzero amplitude for the present state of the universe if, as we claim, the metrics in the sum over histories have no-boundary apart from the surface $S$ at the present time? We do not have a definitive answer,
but one possibility would be if the four dimensional part of the saddle point metric was an inflating universe at early times. Hartle and Hawking [Hartle and Hawking (1983)] have shown that such metrics can be rounded off in the past, without a singular beginning and with curvature bounded well away the Planck value. They give a nonzero value of the no–boundary amplitude for almost any universe that arises from an early period of inflation. Thus to illustrate the top–down approach described above, we consider a simple model with a few positive extrema of the effective potential.

We assume the instability of the inflationary phase can be described as the evolution of a scalar order parameter $\phi$ moving in a double well potential $V(\phi)$. We take the potential to have a broad flat–topped maximum $V_0$ at $\phi = 0$ and a minimum at $\phi_1$. The value at the bottom is the present small cosmological constant $\Lambda$. A concrete example would be gravity coupled to a large number of light matter fields [Starobinsky (1980)]. The trace anomaly generates a potential which has unstable de Sitter space as a self–consistent solution.

We are interested in calculating the no–boundary amplitude of an expanding non–empty region of space–time similar to the one we observe today. In the semi–classical approximation, this will come from one or more saddle points in the action. These correspond to solutions of the Einstein equations. One solution is de Sitter space with the field $\phi$ sitting at the minimum of the potential $V(\phi)$. This will have a very large amplitude, but will be complete empty and therefore does not contribute to the top–down amplitude for a universe like ours. To get an expanding universe with $\Omega_m \sim O(1)$ and with small perturbations that lead to galaxies, it seems necessary to have a period of inflation $^{34}$.

We therefore consider the no–boundary amplitude $^{35} \Phi[\tilde{g}^{3}, K, \phi]$ for a closed inflating universe bounded by a 3–surface $S$ with a large approximately constant Hubble parameter $H = \dot{a}/a$ (and corresponding trace $K = -3\dot{a}/a = -3H$), and a nearly constant field $\phi$ near the top of $V$.

$^{34}$One might think it would be more likely for a universe like ours to arise from a fluctuation of the big de Sitter space directly into a hot Big–Bang, rather than from a homogeneous fluctuation up the potential hill that leads to a period of inflation. The amplitude of a hot Big–Bang fluctuation is much smaller, however, than the amplitude of the inflationary saddle points we discuss below (see also [Albrecht and Sorbo (2004)]). The latter do not directly connect to the large de Sitter space, but they could be connected with very little cost in action by a thin bridge [Hawking (1984b)].

$^{35}$We work in the $K$ representation of the wave function (see e.g., [Hawking (1984b)]), where one replaces $g^{ij}_{3}$ on the three-surface $S$ by $\tilde{g}^{ij}_{3}$, the three-metric up to a conformal factor, and $K$, the trace of the second fundamental form. The action $S_{E}^{k}$ differs from (4.99) in that the surface term has a coefficient $1/3$. 
The value of $\phi$ on $S$ is chosen sufficiently far away from the minimum of $V$ to ensure there are at least enough efoldings of inflation for the universe at the present time to be approximately flat.

We first calculate the wave function for imaginary $K$, or real Euclidean $K_e = iK$, and then analytically continue the result to real Lorentzian $K$. There are two distinct saddle point contributions to the amplitude for an inflating universe in this model [Hawking and Hertog (2002)]. In the first case, the universe is created by the HM instanton with constant $\phi = 0$. Then quantum fluctuations disturb the field, causing it to classically roll down the potential to its prescribed value on $S$. Histories of this kind thus have a long period of inflation, and lead to a perfectly flat universe today. The action of the HM geometry is given by [Hawking and Hertog (2006)]

$$S_{HM}^k(K) = -\frac{12\pi^2}{V_0} \left( 1 - \frac{K_e}{(V_0^2 + K_e^2)^{1/2}} \right),$$

(4.104)

where $K_e = 3b_\tau/b$.

There is, however, a second saddle point contribution which comes from a deformed four sphere, with line element

$$ds^2 = d\tau^2 + b^2(\tau)d\Omega^2_3,$$

(4.105)

where $\phi(\tau)$ varies across the instanton. The Euclidean field equations for $O(4)-$invariant instantons are

$$\phi'' = -K_e \phi' + V_\phi, \quad K_\tau^2 + K_e^2 = -(\phi'_\tau + V),$$

(4.106)

where $\phi' = \phi_{\tau'}$. These equations admit a solution, which is part of a Hawking–Turok instanton [Hawking and Turok (1998)], where $\phi$ slowly rolls up the potential from some value $\phi_0$ at the (regular) South Pole to its prescribed value on the 3–surface $S$. Hence this solution represents a class of histories where the scalar starts as far down the potential as the condition that the present universe be approximately flat allows it to. This naturally leads to fewer efoldings of inflation, and hence a universe that is only approximately flat today. The Euclidean action $S_{HT}^k(K)$ of the deformed four sphere was given in [Hawking and Hertog (2002)], in the approximation that $\phi$ is reasonably small everywhere.

A comparison of the action of both saddle points shows that the deformed four sphere dominates the path integral for amplitudes with real Euclidean $K_e$ on $S$. This would seem to suggest that the universe is least likely to start with $\phi$ at the top of the hill. However, we are interested in the amplitude for an expanding Lorentzian universe, with real Lorentzian $K$ on
S. If one analytically continues the action into the complex $K_e-$plane, one finds the action of the deformed four sphere rapidly increases along the imaginary $K_e-$ axis whereas the real part of $S_{HM}^R$ remains constant, and the dominant contribution to amplitudes for larger $K$ on $S$ actually comes from the HM geometry. The reason for this is that a constant scalar field saves more in gradient energy, than it pays in potential energy for being at the top of the hill. Hence a Lorentzian, expanding universe with large Hubble parameter $H$ is most likely to emerge in an inflationary state, with $\phi$ constant at the maximum of the potential.

Top down cosmology thus predicts that in models like trace anomaly inflation, expanding universes with small perturbations that lead to galaxies, start with a long period of inflation, and are perfectly flat today. Furthermore, as discussed earlier, the precise shape of the primordial fluctuation spectra can be computed from the Euclidean path integral, by perturbatively evaluating around the HM saddle point.

4.3.7.5 Prediction in a Potential Landscape

The predictions we obtained in the previous section extend in a rather obvious way to models where one has a potential landscape. A generic potential landscape admits a large class of alternative inflationary histories with no–boundary initial conditions. There will be HM geometries at all positive saddle points of the potential. For saddle points with more than one descent direction, there will generally be a lower saddle point with only one descent direction, and with lower action. If this descent direction is sharply curved $|V''(0)|/H^2 > 1$, one would not expect a significant top–down amplitude to come from the saddle point. Thus only broad saddle points with a single descent direction will give rise to amplitudes for universes like our own. The requirement that the primordial fluctuations be sufficiently large to form galaxies, however, sets a lower bound on the value of $V_0$ \cite{Hawking and Hertog (2006)}.

Only a few of the saddle points will satisfy the demanding condition that they be broad, because it requires that the scalar field varies by order the Planck value across them. Because the dominant saddle points are in the semi–classical regime, the solutions will evolve from the saddle points to the neighboring minima of $V$. Thus top–down cosmology predicts that only a few of the possible vacua in the landscape will have significant amplitudes.

\footnote{Extra constraints from particle physics, when combined with the cosmological constraints discussed here, will probably further raise the value of $V_0$.}
4.3.7.6 Alternative Proposals

To conclude, we briefly comment on a number of different approaches to the problem of initial conditions in cosmology, and we clarify in what respect they differ from the top–down approach we have put forward.

We have already discussed the pre–Big–Bang cosmologies [Gasperini and Veneziano (1993)], where one specifies initial conditions in the infinite past and follows forward in time a single semi–classical history of the universe. Pre–Big–Bang cosmology is thus based on a bottom–up approach to cosmology. It requires one to postulate a fine–tuned initial state, in order to have a smooth deterministic transition through the Big–Crunch singularity.

We have also discussed the anthropic principle [Barrow and Tipler (1986)]. This can be implemented in top–down cosmology, through the specification of final boundary conditions that select histories where life emerges. Anthropic reasoning within the top–down approach is reasonably well–defined, and useful to the extent that it provides a qualitative understanding for the origin of certain late time conditions that one finds are needed in top–down cosmology.

Eternal Inflation

A different approach to string cosmology has been to invoke the phenomenon of eternal inflation [Vilenkin (1983)] to populate the landscape. There are two different mechanisms to drive eternal inflation, which operate in different moduli space regions of the landscape. In regions where the moduli potential monotonically increases away from its minimum, it is argued that inflation can be sustained forever by quantum fluctuations up the potential hill. Other regions of the landscape are said to be populated by the nucleation of bubbles in meta–stable de Sitter regions. The interior of these bubbles may or may not exit inflation, depending on the shape of the potential across the barrier.

Both mechanisms of eternal inflation lead to a mosaic structure for the universe, where causally disconnected thermalized regions with different values for various effective coupling constants are separated from each other by a variety of inflating patches. It has proven difficult, however, to calculate the probability distributions for the values of the constants that a local observer in an eternally inflating universe would measure. This is because there are typically an infinite number of thermalized regions.

One could also consider the no–boundary amplitude for universes with a mosaic structure. However, these amplitudes would be much lower than
the amplitudes for final states that are homogeneous and lie entirely within a single minimum, because the gradient energy in a mosaic universe contributes positively to the Euclidean action. Histories in which the universe eternally inflates, therefore, hardly contribute to the no–boundary amplitudes we measure. Thus the global structure of the universe that eternal inflation predicts, differs from the global structure predicted by top–down cosmology. Essentially this is because eternal inflation is again based on the classical idea of a unique history of the universe, whereas the top–down approach is based on the quantum sum over histories. The key difference between both cosmologies is that in the proposal based on eternal inflation there is thought to be only one universe with a fractal structure at late times, whereas in top–down cosmology one envisions a set of alternative universes, which are more likely to be homogeneous, but with different values for various effective coupling constants.

It nevertheless remains a challenge to identify predictions that would provide a clear observational discriminant between both proposals. We emphasize, however, that even a precise calculation of conditional probabilities in no–boundary cosmology, which takes in account the back–reaction of quantum fluctuations, will make no reference to the exterior of our past light cone. Indeed, the top–down framework we have put forward indicates that the mosaic structure of an eternally inflating universe is a redundant theoretical construction, which should be excised by Ockham's razor. It appears unlikely, therefore, that something like a volume–weighted probability distribution, which underlies the idea of eternal inflation, can arise from calculations in top–down cosmology. The implementation of selection effects in both approaches is fundamentally different, and this should ultimately translate into distinct predictions for observations.

4.3.7.7 Interpretations

In summary, the bottom up approach to cosmology would be appropriate, if one knew that the universe was set going in a particular way in either the finite or infinite past. However, in the absence of such knowledge one is required to work from the top–down [Hawking and Hertog (2006)].

In a top–down approach one computes amplitudes for alternative histories of the universe with final boundary conditions only. The boundary conditions act as late time constraints on the alternatives and select the subclass of histories that contribute to the amplitude of interest. This enables one to test the proposal, by searching among the conditional probabilities
for predictions of future observations with probabilities near one. In top–
down cosmology the histories of the universe thus depend on the precise
question asked, i.e., on the set of constraints that one imposes. There are
histories in which the universe eternally inflates, or is eleven dimensional,
but we have seen they hardly contribute to the amplitudes we measure.

A central idea that underlies the top–down approach is the interplay
between the fundamental laws of nature and the operation of chance in a
quantum universe. In top–down cosmology, the structure and complexity
of alternative universes in the landscape is predictable from first principles
to some extent, but also determined by the outcome of quantum accidents
over the course of their histories.

We have illustrated our framework in a simple model of gravity coupled
to a scalar with a double well potential, and a small fundamental cosmolog-
cal constant $\Lambda$. Imposing constraints that select the subclass of histories
that are three dimensional and approximately flat at late times, with suffi-
ciently large primordial perturbations for structure formation to occur, we
made several predictions in this model.

In particular we have shown that universes within this class are likely
to emerge in an inflationary state. Furthermore, we were able to identify
the dominant inflationary path as the history where the scalar starts all the
way at the maximum of its potential, leading to a long period of inflation
and a perfectly flat universe today. Moreover, one can calculate the relative
amplitudes of neighboring geometries by perturbatively evaluating the
path integral around the dominant saddle point. Neighboring geometries
 correspond to small quantum fluctuations of various continuous quantities,
like the temperature of the CMB radiation or the expectation values of
 moduli fields. In inflationary universes these fluctuations are amplified and
stretched, generating a pattern of spatial variations on cosmological scales
in those directions of moduli space that are relatively flat. The shape of
these primordial spectra depends on the (no) boundary conditions on the
dominant geometry and provides a strong test of the no–boundary proposal.

When one extends these considerations to a potential that depends on
a multi–dimensional moduli space, one finds that only a few of the minima
of the potential will be populated, i.e., will have significant amplitudes.

The top–down approach we have described leads to a profoundly dif-
ferent view of cosmology, and the relation between cause and effect. Top
down cosmology is a framework in which one essentially traces the histories
backwards, from a space–like surface at the present time. The no–boundary
histories of the universe thus depend on what is being observed, contrary
to the usual idea that the universe has a unique, observer independent history. In some sense no–boundary initial conditions represent a sum over all possible initial states. This is in sharp contrast with the bottom–up approach, where one assumes there is a single history with a well defined starting point and evolution. Our comparison with eternal inflation provides a clear illustration of this. In a cosmology based on eternal inflation there is only one universe with a fractal structure at late times, whereas in top–down cosmology one envisions a set of alternative universes, which are more likely to be homogeneous, but with different values for various effective coupling constants [Hawking and Hertog (2006)].

4.3.8 Brane Cosmology

In this section, mainly following [Brax and Bruck (2003)] and [Langlois (2002)], we give a review of cosmological consequences of the brane world scenario, that is cosmological behavior of a brane–universe, i.e., a 3D space, where ordinary matter is confined, embedded in a higher dimensional space–time.

It has been recently suggested that there might exist some extra spatial dimensions, not in the traditional Kaluza–Klein sense where the extra–dimensions are compactified on a small enough radius to evade detection in the form of Kaluza–Klein modes, but in a setting where the extra dimensions could be large, under the assumption that ordinary matter is confined onto a 3D subspace, called brane (more precisely ‘3–brane’, referring to the three spatial dimensions) embedded in a larger space, called bulk [Langlois (2002)].

Recall that the idea of extra dimensions was proposed in the early twentieth century by Nordstrom and a few years later by Kaluza and Klein [Kaluza (1921)]. It has reemerged over the years in theories combining the principles of quantum mechanics and relativity. In particular theories based on supersymmetry, especially superstring theories, are naturally expressed in more than four dimensions [Polchinski (1998)]. Four dimensional physics is retrieved by Kaluza–Klein reduction, i.e., compactifying on a manifold of small size, typically much smaller than the size of an atomic nucleus.

Recent developments in string theory and its extension M–theory have suggested another approach to compactify extra spatial dimensions. According to these developments, the standard model particles are confined on a brane–hypersurface embedded in a higher dimensional bulk. Only gravity and other exotic matter such as the dilaton can propa-
gate in the bulk. Our universe may be such a brane–like object. This idea was originally motivated phenomenologically (see [Akama (1982); Gibbons and Wiltshire (1982)]) and later revived in string theory. Within the brane world scenario, constraints on the size of extra dimensions become weaker, because the standard model particles propagate only in three spatial dimensions. Newton’s law of gravity, however, is sensitive to the presence of extra–dimensions. Gravity is being tested only on scales larger than a tenth of a millimeter and possible deviations below that scale can be envisaged.

From the string theory point of view, brane worlds of the kind discussed in this review spring from a model suggested by Horava and Witten [Horava and Witten (1996)]. The strong coupling limit of the $E_8 \times E_8$ heterotic string theory at low energy is described by eleven dimensional supergravity with the eleventh dimension compactified on an orbifold with $Z_2$ symmetry, i.e., an interval. The two boundaries of space–time (i.e., the orbifold fixed points) are 10–dimensional planes, on which gauge theories (with the $E_8$ gauge groups) are confined. Later Witten argued that 6 of the 11 dimensions can be consistently compactified on a Calabi–Yau threefold and that the size of the Calabi–Yau manifold can be substantially smaller than the space between the two boundary branes [Witten (1998b)]. Thus, in that limit space–time looks five–dimensional with four dimensional boundary branes [Lukas et al. (1999)]. This provides the underlying picture for many brane world models proposed so far.

Another important ingredient was put forward by Arkani-Hamed, Dimopoulos and Dvali (ADD), [Hamed et al. (1998)], who suggested that by confining the standard model particle on a brane the extra dimensions can be larger than previously anticipated. They considered a flat bulk geometry in $(4 + n)$–dimensions, in which $n$ dimensions are compact with radius $R$ (toroidal topology). The 4D Planck mass $M_P$ and the $(4 + n)$D Planck mass $M_{\text{fund}}$, the gravitational scale of the extra dimensional theory, are related by $M_P^2 = M_{\text{fund}}^2 R^n$. Gravity deviates from Newton’s law only on scales smaller than $R$. Since gravity is tested only down to sizes of around a millimeter, $R$ could be as large as a fraction of a millimeter.

ADD assumed that the bulk geometry is flat. Considerable progress was made by Randall and Sundrum, who considered non–flat, i.e., warped bulk geometries [Randall and Sundrum (1999a); Randall and Sundrum (1999b)]. In their models, the bulk space–time is a slice of Anti–de Sitter space–time, i.e., a space–time with a negative cosmological constant. Their discovery was that, due to the curvature of the bulk space time, Newton’s law
of gravity can be obtained on the brane of positive tension embedded in an infinite extra–dimension. Small corrections to Newton’s law are generated and constrain the possible scales in the model to be smaller than a millimetre.

They also proposed a two–brane model in which the hierarchy problem, i.e., the large discrepancy between the Planck scale at $10^{19}$ GeV and the electroweak scale at 100 GeV, can be addressed. The large hierarchy is due to the highly curved AdS background which implies a large gravitational red shift between energy scale on the two branes. In this scenario, the standard model particles are confined on a brane with negative tension sitting at $y = r_c$, whereas a positive tension brane is located at $y = 0$. The large hierarchy is generated by the appropriate inter–brane distance, i.e., the radion. It can be shown that the Planck mass $M_{Pl}$ measured on the negative tension brane is (for $k = \sqrt{-\Lambda_5 \kappa_5^2 / 6}$) given by $M_{Pl}^2 \approx e^{2kr_c} M_5^3 / k$, where $M_5$ is the 5D Planck mass and $\Lambda_5$ the (negative) cosmological constant in the bulk. Thus, we see that, if $M_5$ is not very far from the electroweak scale $M_W \approx \text{TeV}$, we need $kr_c \approx 50$, in order to generate a large Planck mass on our brane. Hence, by tuning the radius $r_c$ of the extra dimension to a reasonable value, one can get a very large hierarchy between the weak and the Planck scale. Clearly a complete realization of this mechanism requires an explanation for such a value of the radion. In other words, the radion needs to be stabilized at a certain value. The stabilization mechanism is not thoroughly understood, though models with a bulk scalar field have been proposed and have the required properties [Goldberger and Wise (1999)].

Another puzzle which might be addressed with brane models is the cosmological constant problem. One may invoke an extra dimensional origin for the apparent (almost) vanishing of the cosmological constant. The self–tuning idea [Hamed et al. (2000)] advocates that the energy density on our brane does not lead to a large curvature of our universe. On the contrary, the extra dimension becomes highly curved, preserving a flat Minkowski brane with apparent vanishing cosmological constant. Unfortunately, the simplest realization of this mechanism with a bulk scalar field fails due to the presence of a naked singularity in the bulk. This singularity can be shielded by a second brane whose tension has to be fine-tuned with the original brane tension [Forste et al. (2000)]. In a sense, the fine tuning problem of the cosmological constant reappears through the extra dimensional back-door.

Finally, we will later discuss in some detail another spectacular consequence of brane cosmology, namely the possible modification to the Fried-
mann equation at very high energy. This effect was first recognised in [Lukas et al. (2000)] in the context of inflatonary solutions. As we will see, Friedmann’s equation for the Randall–Sundrum model has the form [Cline et al. (1999); Csaki et al. (1999)]

\[ H^2 = \frac{\kappa_5^4}{36} \rho^2 + \frac{8\pi G_N}{3} \rho + \Lambda, \]

relating the expansion rate of the brane \( H \) to the (brane) matter density \( \rho \) and the (effective) cosmological constant \( \Lambda \). The cosmological constant can be tuned to zero by an appropriate choice of the brane tension and bulk cosmological constant, as in the Randall–Sundrum case. Notice that at high energies, for which \( \rho \gg \frac{96\pi G_N}{\kappa_5^4} \), where \( \kappa_5^2 \) is the five-dimensional gravitational constant, the Hubble rate becomes \( H \propto \rho \), while in ordinary cosmology \( H \propto \sqrt{\rho} \). The latter case is retrieved at low energy, i.e., \( \rho \ll \frac{96\pi G_N}{\kappa_5^4} \). Clearly, modifications to the Hubble rate can only be significant before nucleosynthesis. They may have drastic consequences on early universe phenomena such as inflation.

4.3.8.1 The Randall–Sundrum Brane World

Randall–Sundrum Model of the Universe

The so-called Randall–Sundrum (RS) models in particle physics propose that the real world is a higher-dimensional universe described by warped geometry, that is a Lorentzian manifold whose metric tensor can be written in form

\[ ds^2 = g_{ab}(y)dy^ady^b + f(y)g_{ij}dx^idx^j. \]

Note that the geometry almost decomposes into a Cartesian product of the \( y \)-geometry and the \( x \)-geometry, except that the \( x \)-part is warped, i.e., it is rescaled by a scalar function of the other coordinates \( y \). For this reason, the metric of a warped geometry is often called a warped product metric. Warped geometries are useful in that separation of variables can be used when solving PDEs over them.

While studying the so-called technicolor models\(^{37}\) L. Randall and R.  

\(^{37}\)The technicolor models are theories beyond the Standard Model (sometimes, but not always, GUTs) which do not have a scalar Higgs field. Instead, they have a larger number of fermion fields than the Standard Model and involve a larger gauge group. This larger gauge group is spontaneously broken down to the Standard Model group as fermion condensates form. The idea of technicolor is to build a model in which the sort of
Sundrum have proposed in [Randall and Sundrum (1999b)] that our universe is a 5D Anti de Sitter space\textsuperscript{38} and the elementary particles except for the graviton are localized on a (3 + 1)D branes\textsuperscript{39}.

dynamics we see in quantum chromodynamics (QCD) can be used to explain the masses of the W and Z bosons. In QCD, there are quarks that feel both the weak interaction and the strong interaction. The strong interaction binds them together in condensates which spontaneously break electroweak symmetry. In fact, QCD itself gives masses to the W and Z bosons, but these masses are tiny compared to the observed masses.

Technicolor uses a QCD–like theory at a higher energy scale to give the observed masses to the W and Z bosons. Unfortunately the simplest models are already experimentally ruled out by precision tests of the electroweak interactions. There is currently no fully satisfactory model of technicolor, but it is possible that some form of technicolor will be experimentally discovered at the Large Hadron Collider.

Recall that nD anti de Sitter space represents the maximally symmetric, simply–connected, Lorentzian manifold with constant negative curvature. It may be regarded as the Lorentzian analog of nD hyperbolic space. In the language of general relativity, anti de Sitter space is the maximally symmetric, vacuum solution of Einstein’s field equation with a negative cosmological constant $\Lambda$.

A coordinate patch covering part of the space gives the half–space coordinatization of anti de Sitter space. The metric for this patch is

$$ds^2 = \frac{1}{y^2} \left( dt^2 - dy^2 - \sum_i dx_i^2 \right).$$

In the limit as $y = 0$, this reduces to a Minkowski metric

$$dy^2 = \left( dt^2 - \sum_i dx_i^2 \right).$$

Thus, the anti–de Sitter space contains a conformal Minkowski space at infinity, the so–called conformal infinity (‘infinity’ having $y$–coordinate zero in this patch). The constant time slices of this coordinate patch are hyperbolic spaces in the Poincaré half–plane metric.

There are two types of AdS space: one where time is periodic, and the universal cover with non–periodic time. The coordinate patch above covers half of a single period of the space–time. Because the conformal infinity of AdS is time–like, specifying the initial data on a space–like hypersurface would not determine the future evolution uniquely (i.e., deterministically) unless there are boundary conditions associated with the conformal infinity.

Recall that branes or $p$–branes are spatially extended objects that appear in string theory and its relatives (M–theory and brane cosmology). The variable $p$ refers to the spatial dimension of the brane. That is, a 0–brane is a zero–dimensional particle, a 1–brane is a string, a 2–brane is a ‘membrane’, etc. Every $p$–brane sweeps out a $(p + 1)$D world volume as it propagates through space–time.

The so–called brane cosmology refers to several theories in particle physics and cosmology motivated by, but not rigorously derived from, superstring theory and M–theory. The central idea is that our visible, 4D universe is entirely restricted to a brane inside a higher–dimensional space, called the bulk. The additional dimensions may be taken to be compact, in which case the observed universe contains the extra dimensions, and then no reference to the bulk is appropriate in this context. In the bulk model, other
There are two popular RS models. The first, called RSI, has a finite size for the extra dimension with two branes, one at each end. The second, RS2, is similar to the first, but one brane has been placed infinitely far away, so that there is only one brane left in the model.

The RSI model attempts to address the hierarchy problem. The warping of the extra dimension is analogous to the warping of space–time in the vicinity of a massive object, such as a black hole. This warping, or red–shifting, generates a large ratio of energy scales so that the natural energy scale at one end of the extra dimension is much larger than at the other end,

\[
ds^2 = \frac{1}{k^2y^2}(dy^2 + \eta_{\mu\nu} dx^\mu dx^\nu),
\]

where \(k\) is some constant and \(\eta\) has \((-+++)-metric signature. This space has boundaries at \(y = 1/k\) and \(y = 1/Wk\), with \(0 \leq \frac{1}{k} \leq \frac{1}{Wk}\), where \(k\) is around the Planck scale, \(W\) is the warp factor and \(Wk\) is around a TeV. The boundary at \(y = 1/k\) is called the Planck brane and the boundary at \(y = 1/Wk\) is called the TeV–brane. The particles of the Standard Model reside on the TeV brane. The distance between both branes is only \(-ln(W)/k\).

In another coordinate system,

\[
\varphi = -\frac{\pi \ln(ky)}{\ln(W)},
\]

so that \(0 \leq \varphi \leq \pi\) and

\[
ds^2 = \left(\frac{\ln(W)}{\pi k}\right)^2 d\varphi^2 + e^{2\ln(W)x}\eta_{\mu\nu} dx^\mu dx^\nu.
\]

The RSII model uses the same geometry as RSI, but there is no TeV brane. The particles of the standard model are presumed to be on the branes may be moving through this bulk. Interactions with the bulk, and possibly with other branes, can influence our brane and thus introduce effects not seen in more standard cosmological models. As one of its attractive features, the model can explain the weakness of gravity relative to the other fundamental forces of nature, thus solving the so-called hierarchy problem. In the brane picture, the other three forces (electromagnetism and the weak and strong nuclear forces) are localised on the brane, but gravity has no such constraint and so much of its attractive power ‘leaks’ into the bulk. As a consequence, the force of gravity should appear significantly stronger on small (sub-millimetre) scales, where less gravitational force has ‘leaked’. The Randall–Sundrum, pre–Big–Bang, ekpyrotic and cyclic scenarios are particular models of brane cosmology which have attracted a considerable amount of attention. The theory hypothesises that the origin of the Big–Bang could have occurred when two parallel branes touched.
Planck brane. This model was originally of interest because it represented an infinite 5D model which, in many respects, behaved as a 4D model. This setup may also be of interest for studies of the AdS/CFT conjecture.\footnote{The AdS/CFT correspondence \textit{(anti-De-Sitter space/conformal field theory correspondence)}, sometimes called the Maldacena duality, is the conjectured equivalence between a string theory defined on one space, and a quantum field theory without gravity defined on the conformal boundary of this space, whose dimension is lower by at least one. The name suggests that the first space is the product of anti de Sitter space (AdS) with some closed manifold like sphere, orbifold, or noncommutative space, and that the quantum field theory is a conformal field theory (CFT).}

### RS Scenario

Originally, Randall and Sundrum suggested a two–brane scenario in five dimensions with a highly curved bulk geometry as an explanation for the large hierarchy between the Planck scale and the electroweak energy–scale [Randall and Sundrum (1999a)]. In this scenario, the standard model particles live on a brane with (constant) negative tension, whereas the bulk is a slice of Anti–de Sitter (AdS) space–time, i.e., a space–time with a negative cosmological constant. In the bulk there is another brane with positive tension. This is the so–called Randall–Sundrum I (RSI) model. Analysing the solution of Einstein’s equation on the positive tension brane and sending the negative tension brane to infinity, an observer confined to the positive tension brane recovers Newton’s law if the curvature scale of the AdS is smaller than a millimeter [Randall and Sundrum (1999b)]. The higher–dimensional space is non–compact, which must be contrasted with the Kaluza–Klein mechanism, where all extra–dimensional degrees of freedom are compact. This one–brane model, on which we will concentrate here, is the so–called Randall–Sundrum II (RSII) model. It was shown, there is a continuum of Kaluza–Klein modes for the gravitational field, contrasting with the discrete spectrum if the extra dimension is periodic. This leads to a correction to the force between two static masses on the brane. To be specific, it was shown that the potential energy between two point masses confined on the brane is given by

\[
V(r) = G_N m_1 m_2 \frac{r}{r} \left(1 + \frac{l^2}{r^2} + O(r^{-3})\right).
\]

In this equation, \(l\) is related to the 5D bulk cosmological constant \(\Lambda_5\) by

\[
l^2 = -\frac{6}{(\kappa_5^2 \Lambda_5)}
\]

and is therefore a measure of the curvature scale of the bulk space–time. Gravitational experiments show no deviation from Newton’s
law of gravity on length scales larger than a millimeter. Thus, \( l \) has to be smaller than that length scale.

The static solution of the Randall–Sundrum model can be obtained as follows: The total action consists of the Einstein–Hilbert action and the brane action, which in the RS model have the form

\[
S_{\text{EH}} = -\int dx^5 \sqrt{-g^{(5)}} \left( \frac{R}{2\kappa_5^2} + \Lambda_5 \right),
\]

\[
S_{\text{brane}} = \int dx^4 \sqrt{-g^{(4)}} (-\sigma).
\]

The parameter \( \Lambda_5 \) (the bulk cosmological constant) and \( \sigma \) (the brane tension) are constant and \( \kappa_5 \) is the 5D gravitational coupling constant. The brane is located at \( y = 0 \) and we assume a \( Z_2 \) symmetry, i.e., we identify \( y \) with \(-y\). The ansatz for the metric is

\[
ds^2 = e^{-2K(y)} \eta_{\mu\nu} dx^\mu dx^\nu + dy^2.
\]

Einstein’s equations, derived from the action above, give two independent equations:

\[
6K^\prime = -\kappa_5^2 \Lambda_5, \quad 3K'' = \kappa_5^2 \sigma \delta(y).
\]

The first equation can be easily solved:

\[
K = K(y) = \sqrt{-\frac{\kappa_5^2}{6} \Lambda_5} \quad \text{y \equiv ky}, \quad (4.107)
\]

which tells us that \( \Lambda_5 \) must be negative. If we integrate the second equation from \(-\epsilon \) to \(+\epsilon \), take the limit \( \epsilon \to 0 \) and make use of the \( Z_2 \)-symmetry, we get 6\( K'|_0 = \kappa_5^2 \sigma \). Together with (4.107) this tells us that

\[
\Lambda_5 = -\frac{\kappa_5^2}{6} \sigma^2. \quad (4.108)
\]

Thus, there must be a fine-tuning between the brane tension and the bulk cosmological constant for static solutions to exist. Here we will discuss the cosmology of this model in detail.

**Einstein’s Equations on the Brane**

There are two ways of deriving the cosmological equations and we will describe both of them below. The first one is rather simple and makes use of the bulk equations only. The second method uses the geometrical
relationship between 4D and 5D quantities. We begin with the simpler
method.

Friedmann’s Equation from 5D Einstein Equations

In the following subsection we will set \( \kappa_5 \equiv 1 \). We write the bulk metric
as follows:

\[
\text{ds}^2 = a^2 b^2 (dt^2 - dy^2) - a^2 \delta_{ij} dx^i dx^j. \tag{4.109}
\]

This metric is consistent with homogeneity and isotropy on the brane lo-
cated at \( y = 0 \). The functions \( a \) and \( b \) are functions of \( t \) and \( y \) only.
Furthermore, we have assumed flat spatial sections, it is straightforward to
include a spatial curvature. For this metric, Einstein equations in the bulk
read:

\[
a^2 b^2 G_{00} \equiv 3 \left[ \frac{\dot{a}^2}{a^2} + \frac{\dot{a} b}{a b} - \frac{a''}{a} + \frac{a' b'}{a b} + kb^2 \right] = a^2 b^2 \left[ \rho_B + \rho \delta (y - y_b) \right],
\]

\[
a^2 b^2 G_{05} \equiv 3 \left[ \frac{\ddot{a}}{a} - \frac{\dot{a} b}{a b} - 2 \frac{a'^2}{a^2} - \frac{a' b'}{a b} + kb^2 \right] = -a^2 b^2 T_{05}^5,
\]

\[
a^2 b^2 G_{55} \equiv 3 \left[ -\frac{\dot{a}}{a} + 2 \frac{\dot{a} b}{a b} + \frac{\dot{a} b'}{a b} + \frac{b b'}{a b} \right] = -a^2 b^2 T_{55}^5,
\]

\[
a^2 b^2 G_{ij} \equiv 3 \left[ \frac{\ddot{a}}{a} + \frac{\dot{b}}{b} - \frac{\dot{b}^2}{b^2} - 3 \frac{a''}{a} - \frac{b''}{b} + \frac{b'^2}{b^2} + kb^2 \right] \delta_{ij} = -a^2 b^2 \left[ \rho_B + p \delta (y - y_b) \right] \delta_{ij},
\]

where the bulk energy–momentum tensor \( T_{ij}^a \) has been kept general here.
For the RS model we will now take \( \rho_B = -p_B = \Lambda_5 \) and \( T_{00}^i = 0 \). Later we
will use these equations to derive Friedmann’s equation with a bulk scalar
field. In the equations above, a dot represents the derivative with respect
to \( t \) and a prime a derivative with respect to \( y \).

Let us integrate the 00–component over \( y \) from \( -\epsilon \) to \( \epsilon \) and use the fact
that \( a(y) = a(-y), b(y) = b(-y), a'(y) = -a(-y) \) and \( b'(y) = -b(-y) \) (i.e.,
\( \mathbb{Z}_2 \)-symmetry). Then, taking the limit \( \epsilon \to 0 \), we get

\[
\frac{a'}{a} \bigg|_{y=0} = \frac{1}{6} ab \rho. \tag{4.110}
\]
Integrating the $ij$–component in the same way and using the last equation gives

$$\frac{b'}{b} \bigg|_{y=0} = -\frac{1}{2} ab(\rho + p). \quad (4.111)$$

These two conditions are called the junction conditions. The other components of the Einstein equations should be compatible with these conditions.

It is not difficult to show that the restriction of the $05$ component to $y = 0$ leads to

$$\dot{\rho} + 3\frac{\dot{a}}{a} (\rho + p) = 0, \quad (4.112)$$

where we have made use of the junction conditions (4.110) and (4.111). This is nothing but matter conservation on the brane.

Proceeding in the same way with the $55$–component gives

$$\frac{\ddot{a}}{a} - \frac{\dot{a}b}{ab} + kb^2 = -\frac{a^2b^2}{3} \left[ \frac{1}{12} \rho (\rho + 3p) + qB \right].$$

Changing to cosmic time $d\tau = abdt$, writing $a = \exp(\alpha(t))$ and using the energy conservation gives (Flanagan et al. (2000), Bruck et al. (2000))

$$\frac{n(H^2 e^{4\alpha})}{d\alpha} = \frac{2}{3} \Lambda_5 e^{4\alpha} + \frac{n}{d\alpha} \left( e^{4\alpha} \rho^2 \frac{36}{36} \right).$$

In this equation $aH = da/d\tau$. This equation can easily be integrated to give

$$H^2 = \frac{\rho^2}{36} + \frac{\Lambda_5}{6} + \frac{\mu}{a^4}.$$

The final step is to split the total energy–density and pressure into parts coming from matter and brane tension, i.e., to write $\rho = \rho_M + \sigma$ and $p = p_M - \sigma$. Then we find Friedmann’s equation

$$H^2 = \frac{8\pi G}{3} \rho_M \left[ 1 + \frac{\rho_M}{2\sigma} \right] + \frac{\Lambda_4}{3} + \frac{\mu}{a^4}, \quad \text{using} \quad 8\pi G = \frac{\sigma}{18}, \quad \frac{\Lambda_4}{3} = \frac{\sigma^2}{36} + \frac{\Lambda_5}{6}.$$

Comparing the last equation with the fine–tuning relation (4.108) in the static RS solution, we see that $\Lambda_4 = 0$ in this case. If there is a small mismatch between the brane tension and the 5D cosmological constant, then an effective 4D cosmological constant is generated. Another important point is that the 4D Newton constant is directly related to the brane tension. The
constant \( \mu \) appears in the derivation above as an integration constant. The term including \( \mu \) is called the *dark radiation* term [see e.g., \( \text{Mukohyama (2000)} \) \( \text{Ichiki et al. (2002)} \)]. The parameter \( \mu \) can be obtained from a full analysis of the bulk equations [\( \text{Mukohyama et al. (2000)} \)]. An extended version of *Birkhoff’s Theorem* tells us that if the bulk space–time is AdS, this constant is zero [\( \text{Bowcock et al. (2000)} \)]. If the bulk is AdS–Schwarzschild instead, \( \mu \) is non–zero but a measure of the mass of the bulk black hole. In the following we will assume that \( \mu = 0 \) and \( \Lambda_4 = 0 \).

The most important change in Friedmann’s equation compared to the usual 4D form is the appearance of a term proportional to \( \rho^2 \). It tells us that if the matter energy density is much larger than the brane tension, i.e., \( \rho_M \gg \sigma \), the expansion rate \( H \) is proportional \( \rho_M \), instead of \( \sqrt{\rho_M} \). The expansion rate is, in this regime, larger in this brane world scenario. Only in the limit where the brane tension is much larger than the matter energy density, the usual behavior \( H \propto \sqrt{\rho_M} \) is recovered. This is the most important change in brane world scenarios. It is quite generic and not restricted to the Randall–Sundrum brane world model. From Friedmann’s equation and from the energy–conservation equation we can derive *Raychaudhuri’s equation*:

\[
\frac{dH}{d\tau} = -4\pi G (\rho_M + p_M) \left[ 1 + \frac{\rho_M}{\sigma} \right].
\]

We will use these equations later in order to investigate inflation driven by a scalar field confined on the brane.

Notice that at the time of nucleosynthesis the brane world corrections in Friedmann’s equation must be negligible, otherwise the expansion rate is modified and the results for the abundances of the light elements are unacceptably changed. This implies that \( \sigma \geq (1\text{MeV})^4 \). Note, however, that a much stronger constraint arises from current tests for deviation from Newton’s law [\( \text{Maartens (2001)} \)] (assuming the Randall–Sundrum fine–tuning relation (4.108)): \( \kappa_5^3 > 10^5 \text{ TeV} \) and \( \sigma \geq (100\text{GeV})^4 \). Similarly, cosmology constrains the amount of dark radiation. It has been shown that the energy density in dark radiation can at most be 10 percent of the energy density in photons [\( \text{Ida (2000)} \)].

**Another Derivation of Einstein’s Equation**

There is a more powerful way of deriving Einstein’s equation on the brane [\( \text{Shiromizu (2000)} \)]. Consider an arbitrary (3+1)D hypersurface \( \mathcal{M} \) with
unit normal vector $n_a$ embedded in a 5D space–time. The induced metric and the extrinsic curvature of the hypersurface are defined as

$$h_{ab} = \delta_{ab} - n^a n^b, \quad K_{ab} = h^c_a h^b_c \nabla_n n_a.$$  

For the derivation we need three equations, two of them relate 4D quantities constructed from $h_{ab}$ to full 5D quantities constructed from $g_{ab}$. We just state these equations here and refer to [Wald (1984)] for the derivation of these equations. The first equation is the Gauss equation, which reads

$$R^{(4)}_{abcd} = h^c_b h^d_c h^e_a h^f_m R_{f k l m} - 2 K_{[a[c} K_{b]d]}.$$  

This equation relates the 4D curvature tensor $R^{(4)}_{abcd}$, constructed from $h_{ab}$, to the 5D one and $K_{ab}$. The next equation is the Codazzi equation, which relates $K_{ab}$, $n^a$ and the 5D Ricci tensor:

$$\nabla^{(4)}_b K^b_a - \nabla^{(4)}_a K = n^c h^b_c R_{bc}.$$  

One decomposes the 5D curvature tensor $R_{abcd}$ into the Weyl tensor $C_{abcd}$ and the Ricci tensor:

$$R_{abcd} = \frac{2}{3} \left( g_{a[c} R_{n]b} - g_{b[c} R_{n]a} \right) - \frac{1}{6} R g_{a[c} g_{b]n} + C_{abcd}.$$  

If one substitutes the last equation into the Gauss equation and constructs the 4D Einstein tensor, one gets

$$C^{(4)}_{ab} = \frac{2}{3} \left( G_{cd} h^c_a h^b_n + \left( G_{cd} n^c n^a - \frac{1}{4} G \right) h_{ab} \right) + K K_{ab} - K^c_a K_{bc} - \frac{1}{2} \left( K^2 - K_{cd} K_{cd} \right) h_{ab} - E_{ab},$$  

where

$$E_{ab} = C_{acbd} n^c n^d.$$  

We would like to emphasize that this equation holds for any hypersurface. If one considers a hypersurface with energy momentum tensor $T_{ab}$, then there exists a relationship between $K_{ab}$ and $T_{ab}$ ($T$ is the trace of $T_{ab}$) [Isr66):

$$[K_{ab}] = -\kappa_5^2 \left( T_{ab} - \frac{1}{3} h_{ab} T \right),$$  

where [...] denotes the jump:

$$[f](y) = \lim_{\epsilon \to 0} (f(y + \epsilon) - f(y - \epsilon)).$$
These equations are called junction conditions and are equivalent in the cosmological context to the junction conditions (4.110) and (4.111). Splitting $T_{ab} = \tau_{ab} - \sigma h_{ab}$ and inserting the junction condition into equation (4.113), we get Einstein’s equation on the brane:

$$G^{(4)}_{ab} = 8\pi G \tau_{ab} - \Lambda_4 h_{ab} + \kappa^2 \pi_{ab} - E_{ab}. \quad (4.114)$$

The tensor $\pi_{ab}$ is defined as

$$\pi_{ab} = \frac{1}{12} \tau_{ab} - \frac{1}{4} \tau_{ac} \tau^c_b + \frac{1}{8} h_{ab} \tau_{cd} \tau^{cd} - \frac{1}{24} \tau^2 h_{ab},$$

whereas

$$8\pi G = \frac{\kappa^2}{6} \sigma, \quad \Lambda_4 = \frac{\kappa^2}{2} \left( \Lambda_5 + \frac{\kappa^2}{6} \sigma^2 \right).$$

Note that in the Randall–Sundrum case we have $\Lambda_4 = 0$ due to the fine-tuning between the brane tension and the bulk cosmological constant. Moreover $E_{ab} = 0$ as the Weyl–tensor vanishes for an AdS space–time. In general, the energy conservation and the Bianchi identities imply that

$$\kappa^2 \nabla^a \pi_{ab} = \nabla^a E_{ab} \quad (4.115)$$

on the brane.

Clearly, this method is powerful, as it does not assume homogeneity and isotropy nor does it assume that the bulk to be AdS. In the case of an AdS bulk and a Friedmann–Robertson walker brane, the previous equations reduce to the Friedmann equation and Raychaudhuri equation derived earlier. However, the set of equations on the brane are not closed in general [Maartens (2000)], as we will see below.

**Slow–Roll Inflation on the Brane**

We have seen that the Friedmann equation on a brane is drastically modified at high energy where the $\rho^2$ terms dominate. As a result the early universe cosmology on branes tends to be different from standard 4d cosmology. In that vein it seems natural to look for brane effects on early universe phenomena such as inflation [Maartens et al. (2000)] [Copeland et al. (2001)] and on phase–transitions [Brax and Bruck (2003)].

The energy density and the pressure of a scalar field are given by

$$\rho_\phi = \frac{1}{2} \phi_{,\mu} \phi^{,\mu} + V(\phi), \quad p_\phi = \frac{1}{2} \phi_{,\mu} \phi^{,\mu} - V(\phi),$$
where $V(\phi)$ is the potential energy of the scalar field. The full evolution of the scalar field is described by the (modified) Friedmann equation, the Klein–Gordon equation and the Raychaudhuri equation.

We will assume that the field is in a slow–roll regime, the evolution of the fields is governed by (from now on a dot stands for a derivative with respect to cosmic time)

$$3H\dot{\phi} \approx -\frac{\partial V}{\partial \phi}, \quad H^2 \approx \frac{8\pi G}{3} V(\phi) \left( 1 + \frac{V(\phi)}{2\sigma} \right).$$

It is not difficult to show that these equations imply that the slow–roll parameter are given by

$$\epsilon \equiv -\frac{\dot{H}}{H^2} = \frac{1}{16\pi G} \left( \frac{V'}{V} \right)^2 \left[ \frac{4\sigma (\sigma + V)}{2\sigma + V} \right],$$

$$\eta \equiv \frac{V''}{3H^2} = \frac{1}{8\pi G} \left( \frac{V''}{V} \right) \left[ \frac{2\sigma}{2\sigma + V} \right].$$

The modifications to General Relativity are contained in the square brackets of these expressions. They imply that for a given potential and given initial conditions for the scalar field the slow–roll parameters are suppressed compared to the predictions made in General Relativity. In other words, brane world effects ease slow–roll inflation [Maartens et al. (2000)]. In the limit $\sigma \ll V$ the parameter are heavily suppressed. It implies that steeper potentials can be used to drive slow–roll inflation [Copeland et al. (2001)]. Let us discuss the implications for cosmological perturbations.

According to Einstein’s equation (4.114), perturbations in the metric are sourced not only by matter perturbations but also by perturbations of the bulk geometry, encoded in the perturbation of $E_{ab}$. This term can be seen as an external source for perturbations, absent in General Relativity. If one regards $E_{ab}$ as an energy–momentum tensor of an additional fluid (called the Weyl–fluid), its evolution is connected to the energy density of matter on the brane, as one can see from (4.115). If one neglects the anisotropic stress of the Weyl-fluid, then at low energy and superhorizon scales, it decays as radiation, i.e., $\delta E_{ab} \propto a^{-4}$. However, the bulk gravitational field exerts an anisotropic stress onto the brane, whose time–evolution cannot be obtained by considering the projected equations on the brane alone [Maartens (2000)]. Rather, the full 5D equations have to be solved, together with the junction conditions. The full evolution of $E_{ab}$ in the different cosmological eras is currently not understood. However, as we will discuss below, partial results have been obtained for the case of
a de Sitter brane, which suggest that $E_{ab}$ does not change the spectrum of scalar perturbations. It should be noted however, that the issue is not settled and that it is also not clear if the subsequent cosmological evolution during radiation and matter era leaves an imprint of the bulk gravitational field in the anisotropies of the microwave background radiation [Langlois et al. (2001)]. With this in mind, we will, for scalar perturbations, neglect the gravitational backreaction described by the projected Weyl tensor.

Considering scalar perturbations for the moment, the perturbed line element on the brane has the form

$$ds^2 = -(1 + 2A) dt^2 + 2 B dx^i dx^j + ((1 - 2\psi)\delta_{ij} + D_{ij}E) dx^i dx^j,$$

where the functions $A$, $B$, $E$ and $\psi$ depend on $t$ and $x^i$.

An elegant way of discussing scalar perturbations is to make use of of the gauge invariant quantity [Brax and Bruck (2003)]

$$\zeta = \psi + H \frac{\delta \rho}{\dot{\rho}}.$$  (4.118)

In General Relativity, the evolution equation for $\zeta$ can be obtained from the energy–conservation equation. It reads, on large scales [Brax and Bruck (2003)],

$$\dot{\zeta} = -\frac{H}{\rho + p} \delta_{p_{nad}}, \quad \text{where} \quad \delta_{p_{nad}} = \delta_{p_{tot}} - c_s^2 \delta \rho$$  (4.119)

is the non–adiabatic pressure perturbation. The energy conservation equation, however, holds for the Randall–Sundrum model as well. Therefore, (4.119) is still valid for the brane world model we consider. For inflation driven by a single scalar field $\delta p_{nad}$ vanishes and therefore $\zeta$ is constant on superhorizon scales during inflation. Its amplitude is given in terms of the fluctuations in the scalar field on spatially flat hypersurfaces:

$$\zeta = \frac{H \delta \phi}{\dot{\phi}}.$$  (4.120)

The quantum fluctuation in the (slow–rolling) scalar field obey $\langle (\delta \phi)^2 \rangle \approx (H/2\pi)^2$, as the Klein–Gordon equation is not modified in the brane world model we consider. The amplitude of scalar perturbations is [Liddle and Lyth (2000)] $A_S^2 = 4(\zeta^2)/25$. Using the slow–roll equations and (4.120) one
The arguments presented so far suggest that, at least for scalar perturbations, perturbations in the bulk space–time are not important during inflation. This, however, might not be true for tensor perturbations, as gravitational waves can propagate into the bulk. For tensor perturbations, a wave equation for a single variable can be derived (Langlois et al. (2000)). The wave equation can be separated into a 4D and a 5D part, so that the solution has the form $h_{ij} = A(y)h(x^a)e_{ij}$, where $e_{ij}$ is a (constant) polarization tensor. One finds that the amplitude for the zero mode of tensor perturbation is given by (Langlois et al. (2000))

\[ A_T^2 = \frac{4}{25\pi M_{pl}^4} H^2 F^2(H/\mu)|_{k=aH}, \]  

(4.122)

It can be shown that modes with $m > 3H/2$ are generated but they decay during inflation. Thus, one expects in this scenario only the massless modes to survive until the end of inflation (Langlois et al. (2000), Gorbunov et al. (2001)).

From equation (4.122) and (4.121) one sees that the amplitudes of scalar and tensor perturbations are enhanced at high energies. However, scalar perturbations are more enhanced than tensors. Thus, the relative contribution of tensor perturbations will be suppressed, if inflation is driven at high energies.

Finally, we would like to mention that there are also differences between General Relativity and the brane world model we consider for the prediction of two–field brane inflation. Usually correlations are separated in adiabatic and isocurvature modes for two–field inflation. In the RS model, this correlation is suppressed if inflation is driven at high energies. This implies that isocurvature and adiabatic perturbations are uncorrelated, if inflation...
is driven at energies much larger than the brane tension \cite{Brax and Bruck (2003)}.

Coming back to cosmological perturbations, the biggest problem is that the evaluation of the projected Weyl tensor is only possible for the background cosmology. As soon as one tries to analyze the brane cosmological perturbations, one faces the possibility that the $E_{0i}$ terms might not vanish. In particular this means that the equation for the density contrast $\delta = \delta \rho / \rho$, which is given by ($w_m = p/\rho$, $k$ is the wavenumber)

$$\ddot{\delta} + (2 - 3\omega_m)H\dot{\delta} - 6\omega_m(H^2 + \dot{H})\delta = (1 + \omega_m)\delta R_{00} - \omega_m k^2 a^2 \delta,$$

cannot be solved as $\delta R_{00}$ involves $\delta E_{00}$ and can therefore not be deduced solely from the brane dynamics \cite{Maartens (2000)}.

The Randall–Sundrum model discussed here is the simplest brane world model. We have not discussed other important conclusions one can draw from the modifications of Friedmann’s equation, such as the evolution of primordial black holes, its connection to the AdS/CFT correspondence and inflation driven by the trace anomaly of the conformal field theory living on the brane. These developments are important in many respects, because they give not only insights about the early universe but gravity itself.

4.3.8.2 Including a Bulk Scalar Field

Here we are going to generalize the previous results obtained with an empty bulk. To be specific, we will consider the inclusion of a scalar field in the bulk. As we will see, one can extend the projective approach wherein one focuses on the dynamics of the brane, i.e., one studies the projected Einstein and the Klein–Gordon equation \cite{Maeda and Wands (2000), Mennim and Battye (2001)}. As in the RS setting, the dynamics do not closed, as bulk effects do not decouple. We will see that there are now two objects representing the bulk back-reaction: the projected Weyl tensor $E_{\mu\nu}$ and the loss parameter $\Delta \Phi_2$. In the case of homogeneous and isotropic cosmology on the brane, the projected Weyl tensor is determined entirely up to a dark radiation term. Unfortunately, no information on the loss parameter is available. This prevents a rigorous treatment of brane cosmology in the projective approach.

Another route amounts to studying the motion of a brane in a bulk space–time. This approach is successful in the RS case thanks to Birkhoff’s Theorem which dictates a unique form for the metric in the bulk \cite{Bowcock (2000)}.
et al. (2000)]. In the case of a bulk scalar field, no such Theorem is available. One has to resort to various ansatze for particular classes of bulk and brane scalar potentials [Brax and Bruck (2003)].

BPS Backgrounds

Properties of BPS Backgrounds

As the physics of branes with bulk scalar fields is pretty complicated, we will start with a particular example where both the bulk and the brane dynamics are fully under control [Csaki et al. (2000); Youm (2000)]. We specify the bulk Lagrangian as

$$S = \frac{1}{2\kappa_5^2} \int d^5 x \sqrt{-g_5} \left( R - \frac{3}{4} \left( (\partial \phi)^2 + V(\phi) \right) \right),$$

where $V(\phi)$ is the bulk potential. The boundary action depends on a brane potential $U_B(\phi)$

$$S_B = -\frac{3}{2\kappa_5^2} \int d^4 x \sqrt{-g_4} U_B(\phi_0),$$

where $U_B(\phi_0)$ is evaluated on the brane. The BPS backgrounds arise as particular case of this general setting with a particular relationship between the bulk and brane potentials. This relation appears in the study of $N = 2$ supergravity with vector multiplets in the bulk. The bulk potential is given by

$$V = \left( \frac{\partial W}{\partial \phi} \right)^2 - W^2,$$

where $W(\phi)$ is the superpotential. The brane potential is simply given by the superpotential $U_B = W$. We would like to mention, that the last two relations have been also used in order to generate bulk solutions without necessarily imposing supersymmetry. Notice that the RS case can be retrieved by putting $W = \text{cst}$. Supergravity puts further constraints on the superpotential which turns out to be of the exponential type [Brax and Bruck (2003)]

$$W = 4k e^{\alpha \phi},$$
with $\alpha = \frac{1}{\sqrt{12}}, \frac{1}{\sqrt{3}}$. In the following we will often choose this exponential potential with an arbitrary $\alpha$ as an example. The actual value of $\alpha$ does not play any role and will be considered generic.

The bulk equations of motion comprise the Einstein equations and the Klein–Gordon equation. In the BPS case and using the following ansatz for the metric

$$ds^2 = a(y)^2 \eta_{\mu\nu} dx^\mu dx^\nu + dy^2,$$

(4.123)

these second order differential equations reduce to a system of two first order differential equations

$$\frac{a'}{a} = -\frac{W}{4}, \quad \phi' = \frac{\partial W}{\partial \phi}.$$

Notice that when $W = cst$ one easily retrieves the exponential profile of the RS model.

An interesting property of BPS systems can be deduced from the study of the boundary conditions. The Israel junction conditions reduce to

$$\frac{a'}{a}|_B = -\frac{W}{4}|_B$$

and for the scalar field

$$\phi'|_B = \frac{\partial W}{\partial \phi}|_B$$

This is the main property of BPS systems: the boundary conditions coincide with the bulk equations, i.e., as soon as the bulk equations are solved one can locate the BPS branes anywhere in this background, there is no obstruction due to the boundary conditions. In particular two-brane systems with two boundary BPS branes admit moduli corresponding to massless deformations of the background. They are identified with the positions of the branes in the BPS background.

Let us treat the example of the exponential superpotential. The solution for the scale factor reads

$$a = (1 - 4k\alpha^2 x_5)^{1/4\alpha^2},$$

(4.124)

and the scalar field is given by

$$\phi = \frac{-1}{\alpha} \ln(1 - 4k\alpha^2 x_5).$$

(4.125)
For $\alpha \to 0$, the bulk scalar field decouples and these expressions reduce to the RS case. Notice a new feature here, namely the existence of singularities in the bulk, corresponding to

$$a(x_5)|_{x_5} = 0.$$ 

In order to analyse singularities it is convenient to use conformal coordinates: $du = dx_5/a(x_5)$. In these coordinates light follows straight lines $u = \pm t$. It is easy to see that the singularities fall in two categories depending on $\alpha$. For $\alpha^2 < 1/4$ the singularity is at infinity $u_\ast = \infty$. This singularity is null and absorbs incoming gravitons. For $\alpha^2 > 1/4$ the singularity is at finite distance. It is time-like and not wave-regular, i.e., the propagation of wave packets is not uniquely defined in the vicinity of the singularity. For all these reasons these naked singularities in the bulk are a major drawback of brane models with bulk scalar fields [Brax and Bruck (2003)]. In the two-brane case the second brane has to sit in front of the naked singularity.

**De Sitter and Anti de Sitter Branes**

Let us modify slightly the BPS setting by detuning the tension of the BPS brane. This corresponds to adding or subtracting some tension compared to the BPS case, $U_B = TW$, where $T$ is real number. Notice that this modification only affects the boundary conditions, the bulk geometry and scalar field are still solutions of the BPS equations of motion. In this sort of situation, one can show that the brane does not stay static. For the detuned case, the result is either a boosted brane or a rotated brane. We will soon generalize these results so we postpone the detailed explanation to later. Defining by $u(x^\mu)$ the position of the brane in conformal coordinates, one gets

$$(\partial u)^2 = \frac{1 - T^2}{T^2}.$$ 

The brane velocity vector $\partial_\mu u$ is of constant norm. For $T > 1$, the brane velocity vector is time-like and the brane moves at constant speed. For $T < 1$ the brane velocity vector is space-like and the brane is rotated. Going back to a static brane, we see that the bulk geometry and scalar field become $x^\mu$ dependent.

Let us consider the brane geometry when $T > 1$. In particular one can
study the Friedmann equation for the induced bulk factor

\[ H^2 = \frac{T^2 - \frac{1}{16}}{W^2}, \]

where \( W \) is evaluated on the brane. Of course we get the fact that cosmological solutions are only valid for \( T > 1 \). Now in the RS case \( W = 4k \) leading to

\[ H^2 = (T^2 - 1)k^2. \]

In the case \( T > 1 \) the brane geometry is driven by a positive cosmological constant. This is a de Sitter brane. When \( T < 1 \) the cosmological constant is negative, corresponding to an AdS brane. We are going to generalize these results by considering the projective approach to the brane dynamics.

**Bulk Scalar Fields and the Projective Approach**

**The Friedmann Equation**

We will first follow the traditional coordinate dependent path. This will allow us to derive the matter conservation equation, the Klein–Gordon and the Friedmann equations on the brane. Then we will concentrate on the more geometric formulation where the role of the projected Weyl tensor will become transparent \[ [\text{Brax et al. (2001)}][\text{Brax and Bruck (2003)}]. \] Again, in this subsection we will put \( \kappa_5 \equiv 1 \).

We consider a static brane that we choose to put at the origin \( x_5 = 0 \) and impose \( b(0, t) = 1 \). This guarantees that the brane and bulk expansion rates

\[ 4H = \partial_\tau \sqrt{-g_{\mu\nu}}|_0, \quad 3H_B = \partial_\tau \sqrt{-g_4}|_0 \]

coincide. We have identified the brane cosmic time \( d\tau = ab|_0 dt \). We will denote by prime the normal derivative \( \partial_n = \frac{1}{ab}|_0 \partial_x^5 \). Moreover we now allow for some matter to be present on the brane

\[ \tau^{\mu}_{\nu} \text{matter} = (-\rho_m, p_m, p_m, p_m). \]

The bulk energy-momentum tensor reads

\[ T_{ab} = \frac{3}{4} (\partial_a \phi \partial_b \phi) - \frac{3}{8} g_{ab} \left( (\partial \phi)^2 + V \right). \]
The total matter density and pressure on the brane are given by

\[ \rho = \rho_m + \frac{3}{2} U_B, \quad p = p_m - \frac{3}{2} U_B. \]

The boundary condition for the scalar field is unchanged by the presence of matter on the brane.

The (05) Einstein equation leads to matter conservation

\[ \dot{\rho}_m = -3H(\rho_m + p_m). \]

By restricting the (55) component of the Einstein equations we get

\[ H^2 = \frac{\rho^2}{36} - \frac{2}{3} Q - \frac{1}{9} E + \frac{\mu}{a^4} \]

in units of \( \kappa^2 \). The last term is the dark radiation, whose origin is similar to the RS case. The quantity \( Q \) and \( E \) satisfy the differential equations [Bruck et al. (2000)]

\[ \dot{Q} + 4HQ = HT_0^5, \quad \dot{E} + 4HE = -\rho T_0^5. \]

These equations can be easily integrated to give

\[ H^2 = \frac{\rho_m^2}{36} + \frac{U_B \rho_m}{12} - \frac{1}{16a^4} \int d\tau \frac{da^4}{d\tau} (\dot{\phi}^2 - 2U) - \frac{1}{12a^4} \int d\tau a^4 \rho_m \frac{dU_B}{d\tau}, \]

up to a dark radiation term and we have identified

\[ U = \frac{1}{2} \left( U_B^2 - \left( \frac{\partial U_B}{\partial \phi} \right)^2 + V \right). \]

This is the Friedmann equation for a brane coupled to a bulk scalar field. Notice that retarded effects springing from the whole history of the brane and scalar field dynamics are present. Below we will see that these retarded effects come from the projected Weyl tensor. They result from the exchange between the brane and the bulk. Notice, that Newton’s constant depends on the value of the bulk scalar field evaluated on the brane \( \phi_0 = \phi(t, y = 0) \):

\[ \frac{8\pi G_N(\phi_0)}{3} = \kappa^2 U_B(\phi_0). \]

On cosmological scale, time variation of the scalar field induce a time variations of Newton’s constant. This is highly constrained experimentally, leading to tight restrictions on the time dependence of the scalar field [Brax and Bruck (2003)].
To get a feeling of the physics involved in the Friedmann equation, it is convenient to assume that the scalar field is evolving slowly on the scale of the variation of the scale factor. Neglecting the evolution of Newton’s constant, the Friedmann equation reduces to

\[ H^2 = \frac{8\pi G_N(\phi)}{3} \rho_m + U - \frac{\dot{\phi}^2}{16}. \]

Several comments are in order. First of all we have neglected the contribution due to the \( \rho_m^2 \) term as we are considering energy scales below the brane tension. Then the main effect of the scalar field dynamics is to involve the potential energy \( U \) and the kinetic energy \( \dot{\phi}^2 \). Although the potential energy appears with a positive sign we find that the kinetic energy has a negative sign. For an observer on the brane this looks like a violation of unitarity. The minus sign for the kinetic energy is due to the fact that one does not work in the Einstein frame where Newton’s constant does not vary, a similar minus sign appears also in the effective 4D theory when working in the brane frame.

The time dependence of the scalar field is determined by the Klein–Gordon equation. The dynamics is completely specified by

\[ \ddot{\phi} + 4H\dot{\phi} + \left( \frac{1}{3} - \omega_m \right) \rho_m \frac{\partial U_B}{\partial \phi} = -\frac{\partial U}{\partial \phi} + \Delta \Phi_2, \]

where \( p_m = \omega_m \rho_m \). We have identified

\[ \Delta \Phi_2 = \phi''_0 - \frac{\partial U_B}{\partial \phi} \frac{\partial^2 U_B}{\partial \phi^2} |_0. \]

This cannot be set to zero and requires the knowledge of the scalar field in the vicinity of the brane. When we discuss cosmological solutions below, we will assume that this term is negligible.

The evolution of the scalar field is driven by two effects. First of all, the scalar field couples to the trace of the energy momentum tensor via the gradient of \( U_B \). Secondly, the field is driven by the gradient of the potential \( U \), which might not necessarily vanish.

*The Friedmann Equation vs the Projected Weyl Tensor*

We are now coming back to the origin of the non–trivial Friedmann equation. Using the *Gauss–Codazzi equation* one can get the Einstein equation
Quantum Leap

\[ G_{ab} = \frac{3}{8} U h_{ab} + \frac{U_B}{4} \tau_{ab} + \pi_{ab} + \frac{1}{2} \partial_a \phi \partial_b \phi - \frac{5}{16} (\partial \phi)^2 h_{ab} - E_{ab}. \]

Now the projected Weyl tensor can be determined in the homogeneous and isotropic cosmology case. Indeed only the \( E_{00} \) component is independent. Using the Bianchi identity \( \bar{D}^a \bar{G}_{ab} = 0 \) where \( \bar{D}_a \) is the brane covariant derivative, one gets that

\[ \dot{E}_{00} + 4H E_{00} = \partial_\tau \left( \frac{3}{16} \dot{\phi}^2 + \frac{3}{8} U \right) + \frac{3}{2} h^{\phi^2} + \frac{\dot{U}_B}{4} \rho_m, \]

leading to

\[ E_{00} = \frac{1}{a^4} \int d\tau a^4 \left( \partial_\tau \left( \frac{3}{16} \dot{\phi}^2 + \frac{3}{8} U \right) + \frac{3}{2} h^{\phi^2} + \frac{\dot{U}_B}{4} \rho_m \right). \]

Upon using

\[ \bar{G}_{00} = 3H^2, \]

one gets the Friedmann equation. It is remarkable that the retarded effects in the Friedmann equation all spring from the projected Weyl tensor. Hence the projected Weyl tensor proves to be much richer in the case of a bulk scalar field than in the empty bulk case.

**Self–Tuning and Accelerated Cosmology**

Dynamics of the brane is not closed, it is an open system continuously exchanging energy with the bulk. This exchange is characterized by the dark radiation term and the loss parameter. Both require a detailed knowledge of the bulk dynamics. This is of course beyond the projective approach where only quantities on the brane are evaluated. In the following we will assume that the dark radiation term is absent and that the loss parameter is negligible. Furthermore, we will be interested in the effects of a bulk scalar field for late–time cosmology (i.e., well after nucleosynthesis) and not in the case for inflation driven by a bulk scalar field [Brax and Bruck (2003)].

Let us consider the self–tuned scenario as a solution to the cosmological constant problem. It corresponds to the BPS superpotential with \( \alpha = 1 \). In that case the potential \( U = 0 \) for any value of the brane tension. The potential \( U = 0 \) can be interpreted as a vanishing of the brane cosmological constant. The physical interpretation of the vanishing of the cosmological constant is that the brane tension curves the fifth dimensional space–time.
leaving a flat brane intact. Unfortunately, the description of the bulk geometry in that case has shown that there was a bulk singularity which needs to be hidden by a second brane whose tension is fine-tuned with the first brane tension. This reintroduces a fine-tuning in the putative solution to the cosmological constant problem [Forste et al. (2000)].

Let us generalize the self-tuned case to $\alpha \neq 1$, i.e., $U_B = TW$, $T > 1$ and $W$ is the exponential superpotential. The resulting induced metric on the brane is of the FRW type with a scale factor

$$a(t) = a_0 \left( \frac{t}{t_0} \right)^{1/3 + 1/6\alpha^2},$$

leading to an acceleration parameter

$$q_0 = \frac{6\alpha^2}{1 + 2\alpha^2} - 1.$$ 

For the supergravity value $\alpha = -\frac{1}{\sqrt{12}}$ this leads to $q_0 = -4/7$. This is in coincidental agreement with the supernovae results. This model can serve as a brane quintessence model [Brax and Bruck (2003)].

The Brane Cosmological Eras

Let us now consider the possible cosmological scenarios with a bulk scalar field [Brax et al. (2001); Brax and Bruck (2003)]. We assume that the potential energy of the scalar field $U$ is negligible throughout the radiation and matter eras before serving as quintessence in the recent past.

Assuming radiation domination, the scale factor behaves like

$$a = a_0 \left( \frac{t}{t_0} \right)^{1/4},$$

and the scalar field

$$\phi = \phi_i + \beta \ln \left( \frac{t}{t_0} \right).$$

In the radiation dominated era, no modification is present, provided $\phi = \phi_i$, which is a solution of the Klein–Gordon equation as the trace of the energy–momentum of radiation vanishes (together with a decaying solution, which we have neglected). In the matter dominated era the scalar field evolves due to the coupling to the trace of the energy–momentum tensor.
has two consequences. Firstly, the kinetic energy of the scalar field starts contributing in the Friedmann equation. Secondly, the effective Newton constant does not remain constant. The cosmological evolution of Newton’s constant is severely constrained since nucleosynthesis. This restricts the possible time variation of $\phi$ [Brax and Bruck (2003)].

In order to be more quantitative let us come back to the exponential superpotential case with a detuning parameter $T$. The time dependence of the scalar field and scale factor become

$$\phi = \phi_1 - \frac{8}{15} \alpha \ln \left( \frac{t}{t_e} \right), \quad a = a_e \left( \frac{t}{t_e} \right)^{\frac{2}{3} - \frac{8}{5} \alpha^2},$$

where $t_e$ and $a_e$ are the time and scale factor at matter-radiation equality. Notice the slight discrepancy of the scale factor exponent with the standard model value of $2/3$. The redshift dependence of the Newton constant is

$$\frac{G_N(z)}{G_N(z_e)} = \left( \frac{z + 1}{z_e + 1} \right)^{4\alpha^2/5}.$$

For the supergravity model with $\alpha = -\frac{1}{\sqrt{12}}$ and $z_e \sim 10^3$ this leads to a decrease by (roughly) $37\%$ since nucleosynthesis. This is marginally compatible with experiments [Brax and Bruck (2003)].

Finally let us analyse the possibility of using the brane potential energy of the scalar field $U$ as the source of acceleration now. We have seen that when matter is negligible on the brane, one can build brane quintessence models. We now require that this occurs only in the recent past. As can be expected, this leads to a fine-tuning problem as $M^4 \sim \rho_c$, where $M^4 = (T - 1) \frac{3W}{2\kappa^2}$ is the amount of detuned tension on the brane. Of course this is nothing but a reformulation of the usual cosmological constant problem. Provided one accepts this fine–tuning, as in most quintessence models, the exponential model with $\alpha = -\frac{1}{\sqrt{12}}$ is a cosmological consistent quintessence model with a five dimensional origin.

Therefore, the main difference between a brane world model with a bulk scalar field and the Randall–Sundrum model is that the gravitational constant becomes time–dependent. As such it has much in common with scalar–tensor theories [Fujii and Maeda (2003)], but there are important differences due to the projected Weyl tensor $E_{\mu\nu}$ and its time–evolution. The bulk scalar field can play the role of the quintessence field, as discussed above, but it could also play a role in an inflationary era in the very early universe. In any case, the cosmology of such a system is much richer and,
because of the variation of the gravitational constant, more constrained. It
remains to be seen if the bulk scalar field can leave a trace in the CMB
anisotropies and Large Scale Structures (for first results see [Brax et al.
(2001)].)

4.3.8.3 Moving Branes in a Static Bulk
So far, we were mostly concerned with the evolution of the brane, without
referring to the bulk itself. In fact, the coordinates introduced in (4.109)
are a convenient choice for studying the brane itself, but when it comes to
analysing the bulk dynamics and its geometry, these coordinates are not the
best choice. We have already mentioned the extended Birkhoff Theorem
above. It states that for the case of a vacuum bulk space–time, the bulk is
necessarily static, in certain coordinates. A cosmological evolving brane is
then moving in that space–time, whereas for an observer confined on the
brane the motion of the brane will be seen as an expanding (or contracting)
universe. In the case of a scalar field in the bulk, a similar Theorem is
unfortunately not available, which makes the study of such systems much
more complicated. We will now discuss these issues in some detail, following
in particular [Davis (2002a)] and [Davis (2002b)].

Motion in AdS–Schwarzschild Bulk

We have already discussed the static background associated with BPS
configurations (including the Randall–Sundrum case) above. Here we will
focus on other backgrounds for which one can integrate the bulk equations
of motion. Let us write the following ansatz for the metric
\[ ds^2 = -A^2(r)dt^2 + B^2(r)dr^2 + R^2(r)d\Sigma^2, \]
where \( d\Sigma^2 \) is the metric on the 3d symmetric space of curvature \( q = 0, \pm 1 \).
In general, the function \( A, B \) and \( R \) depend on the type of scalar field potential.
This is to be contrasted with the case of a negative bulk cosmological constant
where Birkhoff’s Theorem states that the most general solution of the (bulk)
Einstein equations is given by \( A^2 = f, B^2 = 1/f \) and \( R = r \)
where \( f(r) = q + r^2/l^2 - \mu/r^2 \). We have denoted by \( l = 1/k = \sqrt{-6/(\Lambda_5 \kappa_5^2)} \)
the AdS scale and \( \mu \) the black hole mass. This solution is the so–called
AdS–Schwarzschild solution.

Let us now study the motion of a brane of tension \( T/l \) in such a back-
ground. The equation of motion is determined by the junction conditions.
Quantum Leap

The method will be reviewed later when a scalar field is present in the bulk. The resulting equation of motion for a boundary brane with a $Z_2$ symmetry is

$$(\dot{r}^2 + f(r))^{1/2} = \frac{T}{r},$$

for a brane located at $r$. Here $\dot{r}$ is the velocity of the brane measured with the proper time on the brane. This leads to the following Friedmann equation

$$H^2 \equiv \left(\frac{\dot{r}}{r}\right)^2 = \frac{T^2 - 1}{l^2} - \frac{q}{r^2} + \frac{\mu}{r^4}.$$

So the brane tension leads to an effective cosmological constant $(T^2 - 1)/l^2$. The curvature gives the usual term familiar from standard cosmology while the last term is the dark radiation term whose origin springs from the presence of a black–hole in the bulk. At late time the dark radiation term is negligible for an expanding universe, we retrieve the cosmology of a FRW universe with a non–vanishing cosmological constant. The case $T = 1$ corresponds of course to the RS case.

Moving Branes

Let us now describe the general formalism, which covers the case of the AdS–Schwarzschild space–time mentioned above. Consider a brane embedded in a static background. It is parametrized by the coordinates $X^A(x^\mu)$ where $A = 0 \ldots 4$ and the $x^\mu$ are world volume coordinates. Locally the brane is characterized by the local frame

$$e^A_\mu = \frac{\partial X^A}{\partial x^\mu},$$

which are tangent to the brane. The induced metric is given by

$$h_{\mu\nu} = g_{AB}e^A_\mu e^B_\nu,$$

and the extrinsic curvature

$$K_{\mu\nu} = e^A_\mu e^B_\nu D_A n_B,$$

where $n^A$ is the unit vector normal to the brane defined by (up to a sign ambiguity)

$$g_{AB}n^A n^B = 1, \quad n_A e^A_\mu = 0.$$
For a homogeneous brane embedded in the space–time described by the metric (4.126), we have $T = T(\tau), \ r = r(\tau)$ where $\tau$ is the proper time on the brane. The induced metric is

$$ds^2_B = -d\tau^2 + R^2(\tau)d\Sigma^2.$$ 

The local frame becomes

$$e^A_{\tau} = (\dot{T}, \dot{r}, 0, 0, 0), \quad e^A_i = (0, 0, \delta^A_i),$$

while the normal vector reads

$$n_A = (AB\dot{r}, -B\sqrt{1 + \dot{r}^2}, 0, 0, 0).$$

The components of the extrinsic curvature tensor can found to be

$$K_{ij} = -\sqrt{1 + B^2\dot{r}^2}B\delta_{ij}, \quad K_{\tau\tau} = \frac{1}{ABd\tau}(A\sqrt{1 + B^2\dot{r}^2}).$$

The junction conditions are given by

$$K_{\mu\nu} = -\frac{\kappa^2}{2} \left( \tau_{\mu\nu} - \frac{1}{3} \tau h_{\mu\nu} \right).$$

This implies that the brane dynamics are specified by the equations of motion

$$\sqrt{1 + B^2\dot{r}^2} \frac{R'}{R} = \frac{\kappa^2}{6} \rho, \quad \text{and}$$

$$\frac{1}{ABd\tau} \left( A\sqrt{1 + B^2\dot{r}^2} \right) = -\frac{\kappa^2}{6} (2\rho + 3p),$$

where we have assumed a fluid description for the matter on the brane. These two equations determine the dynamics of any brane in a static background.

Let us now close the system of equations by stating the scalar field boundary condition [Brax and Bruck (2003)]

$$n^A \partial_A \phi = \frac{\kappa^2}{2} \frac{d\xi}{d\phi} (\rho - 3p),$$

where the coupling to the brane is defined by the Lagrangian

$$S_{brane} = \int d^4x \mathcal{L}[\psi_m, \bar{h}_{\mu\nu}].$$
where $\psi_m$ represents the matter fields and
$$\tilde{h}_{\mu\nu} = e^{2\xi(\phi)}h_{\mu\nu}.$$ This reduces to
$$\phi' = \frac{\kappa^2}{2} \frac{B}{\sqrt{1 + B^2r^2}} \frac{d\xi}{d\phi} (-\rho + 3p).$$ Combining the junction conditions leads to the conservation equation
$$\dot{\rho} + 3H(\rho + p) = (\rho - 3p)\dot{\xi}.$$ This is nothing but the conservation of matter in the Jordan frame defined by $\tilde{h}_{\mu\nu}$.

We now turn to a general analysis of the brane motion in a static bulk. To do that it is convenient to parametrize the bulk metric slightly differently
$$ds^2 = -f^2(r)h(r)dt^2 + \frac{dr^2}{h(r)} + r^2d\Sigma^2.$$ Now, the Einstein equations lead to (redefining $\phi \rightarrow \frac{\sqrt{3}}{2\kappa^2} \phi$ and $V \rightarrow \frac{3}{8\kappa^2} V$)

$$\frac{3}{r^2\kappa^2} \left( h + \frac{rh'}{2} - q \right) = -\kappa^2 \left( \frac{h\phi'^2}{2} + V \right), \quad (4.127)$$

$$\frac{3}{r^2\kappa^2} \left( h + \frac{rh'}{2} - q + \frac{hf'}{f} \right) = \kappa^2 \left( \frac{h\phi'^2}{2} - V \right), \quad (4.128)$$

and the Klein–Gordon equation
$$h\phi'' + \left( \frac{3h}{r} + \frac{hf'}{f} + h' \right) \phi' = \frac{dV}{d\phi}.$$ Subtracting (4.127) from (4.128) and solving the resulting differential equation, we get
$$f = \exp \left( \frac{\kappa^2}{3} \int dr' r'^2 \phi'^2 \right).$$ It is convenient to evaluate the spatial trace of the projected Weyl tensor. This is obtained by computing both the bulk Weyl tensor and the vector normal to the moving brane. With $A = \sqrt{h} f$, $R = r$, $B = 1/\sqrt{h}$, this gives
$$\frac{\mu}{r^4} = -\frac{E_1}{3} = \frac{r}{4f^2} \left( \frac{hf'^2}{r^2} \right)' + \frac{q}{2f^2}.$$
This is the analogue of the dark radiation term for a general background. The equations of motion can be cast in the form

\[ \mu' = -\frac{\kappa_5^2}{3}(\mu - \frac{kr^2}{2})r \phi'^2, \]
\[ \dot{\mathcal{H}} + 4\frac{\mu}{r} = -\frac{2\kappa_5^2}{3}(\mathcal{H} - \frac{q}{r^2})r \phi'^2, \]
\[ \kappa_5^2 V = 6\mathcal{H} + \frac{3}{4}r \dot{\mathcal{H}} - 3\frac{\mu}{r^2}, \quad \text{using} \quad \mathcal{H} = \frac{q - h}{r^2}. \]

This allows to retrieve easily some of the previous solutions. Choosing \( \phi \) to be constant leads to \( f = 1, \mu \) is constant and

\[ \mathcal{H} = -\frac{1}{r^2} + \frac{\mu}{r^3}. \]

This is the AdS–Schwarzschild solution.

For \( q = 0 \) the equations of motion simplify to

\[ \frac{\kappa_5^2}{3} r \phi' = -\frac{n \ln \mu}{d\phi}, \quad (4.129) \]
\[ n \left( \frac{\mathcal{H}}{\mu^2} \right) = \frac{1}{\mu} n \left( r^{-4} \right), \]
\[ \frac{\kappa_5^2}{6} V = -\frac{3}{4\kappa_5^2} \frac{d\mu}{d\phi} n \left( \frac{\mathcal{H}}{\mu} \right) + \mathcal{H} \quad (4.130) \]

In this form it is easy to see that the dynamics of the bulk are completely integrable. First of all the solutions depend on an arbitrary function \( \mu(\phi) \) which determines the dynamics. Notice that \( f = \mu_0/\mu \), where \( \mu_0 \) is an arbitrary constant. The radial coordinate \( r \) is obtained by simple integration of (4.129)

\[ r = r_0 e^{-r^2/2} \int \frac{d\phi}{\mu^2} \int \frac{d\phi}{\mu}. \]

Finally the rest of the metric follows from

\[ h = -\frac{4\kappa_5^2}{3} r^2 \mu^2 \int d\phi \frac{d\phi}{d\mu} e^{\frac{1}{4} \int \frac{d\phi}{\mu^2}}. \]

The potential \( V \) then follows [4.130]. This is remarkable and shows why Birkhoff’s Theorem is not valid in the presence of a bulk scalar field. Moreover, it is intriguing that the generalization of the dark energy term dictates the bulk dynamics completely.
It is interesting to recast the Friedmann equation in the form

\[ H^2 = \mathcal{H} + \frac{\kappa_5^4}{36} \mu^2 \rho^2, \]

where \( H \) is the Hubble parameter on the brane in cosmic time. One can retrieve standard cosmology by studying the dynamics in the vicinity of a critical point \( \frac{d\mu}{d\phi} = 0 \). Parameterizing

\[ \mu = \frac{6A}{\kappa_5^2} + B\phi^2, \]

leads to the Friedmann equation

\[ H^2 = \frac{\kappa_5^4}{36} (\rho^2 - \theta)\mu^2 + \frac{\mu}{a^4} + o(a^{-4}). \]

Here \( \theta \) is an arbitrary integration constant. Notice that this is a small deviation from the RS case as \( \phi = r^{-B/A} \) goes to zero at large distances. Hence, standard cosmology is retrieved at low energy and long distance.

4.3.8.4 Cosmology of a Two–Brane System

Here we will once more include an ingredient suggested by particle physics theories, in particular M–theory. So far we have assumed that there is only one brane in the whole space–time. According to string theory, there should be at least another brane in the bulk. Indeed, in heterotic M–theory these branes are the boundaries of the bulk space–time [Horava and Witten (1996)]. However, even from a purely phenomenological point of view there is a reason to include a second brane: the bulk singularity (or the AdS horizon). As we have seen above, the inclusion of a bulk scalar field often implies the presence of a naked singularity located away from the positive tension brane. The second brane which we include now should shield this singularity, so that the physical space–time stretches between the two branes. Another motivation is the hierarchy problem. Randall and Sundrum proposed a two brane model (one with positive and one with negative tension), embedded in a 5D AdS space–time. In their scenario the standard model particles would be confined on the \emph{negative tension brane}. As they have shown, in this case gravity is weak due to the warping of the bulk space–time. However, as will become clear from the above results, in order for this model to be consistent with gravitational experiments, the interbrane distance has to be fixed [Garriga and Tanaka (2000)]. This can
be achieved, for example, with a bulk scalar field. As shown in [Garriga and Tanaka (2000)], gravity in the two–brane model of RS is described by a scalar–tensor theory, in which the interbrane–distance, called radion, plays the role of a scalar field. The bulk scalar field will modify the Brans–Dicke parameter of the scalar field and will introduce a second scalar field in the low–energy effective theory, so that the resulting theory at low energy in the case of two branes and a bulk scalar field is a bi–scalar–tensor theory [Brax and Bruck (2003)]. In the following we will investigate the cosmological consequences when the distance between the branes is not fixed (for some aspects not covered here see e.g., [Csaki et al. (2000); Langlois and Sorbo (2002)]).

The Low–Energy Effective Action

In order to understand the cosmology of the two–brane system, we derive the low-energy effective action by utilizing the moduli space approximation. From the above discussion, it becomes clear, that the general solution of the bulk Einstein equations for a given potential is difficult to find. The moduli space approximation gives the low–energy–limit effective action for the two brane system, i.e., for energies much smaller than the brane tensions.

In the static BPS solutions described above, the brane positions can be chosen arbitrarily. In other words, they are moduli fields. It is expected that by putting some matter on the branes, these moduli field become time-dependent, or, if the matter is inhomogeneously distributed, space–time dependent. Thus, the first approximation is to replace the brane–positions with space–time dependent functions. Furthermore, in order to allow for the gravitational zero–mode, we will replace the flat space–time metric $\eta_{\mu\nu}$ with $g_{\mu\nu}(x^\alpha)$. We do assume that the evolution of these fields is slow, which means that we neglect terms like $(\partial\phi)^3$ when constructing the low-energy effective action.

As already mentioned, the moduli space approximation is only a good approximation at energies much less than the brane tension. Thus, we do not recover the quadratic term in the moduli space approximation. We are interested in the late time effects after nucleosynthesis, where the corrections have to be small.

Replacing $\eta_{\mu\nu}$ with $g_{\mu\nu}(x^\alpha)$ in (4.123) and collecting all the terms one
finds from the 5D action after an integration over \( y \):

\[
S_{\text{MSA}} = \int d^4x \sqrt{-g_4} \left[ f(\phi, \sigma) R^{(4)} + \frac{3}{4} a^2(\phi) \frac{U_B(\phi)}{\kappa_5^2} (\partial \phi)^2 \right. \\
- \left. \frac{3}{4} a^2(\sigma) \frac{U_B(\sigma)}{\kappa_5^2} (\partial \sigma)^2 \right], \\
\text{with} \\
f(\phi, \sigma) = \frac{1}{\kappa_5^2} \int_\phi^\sigma dy a^2(y),
\]

and \( a(y) \) given by (4.124). The moduli \( \phi \) and \( \sigma \) represent the location of the two branes. Note that the kinetic term of the field \( \phi \) has the wrong sign. This is an artifact of the frame we use here. As we will see below, it is possible to go to the Einstein frame with a simple conformal transformation, in which the sign in front of the kinetic term is correct for both fields.

In the following we will concentrate on the BPS system with above exponential superpotential. Let us redefine the fields according to

\[
\tilde{\phi}^2 = (1 - 4k\alpha^2 \phi)^{2\beta}, \quad \tilde{\sigma}^2 = (1 - 4k\alpha^2 \sigma)^{2\beta},
\]

(4.131)

with \( \beta = \frac{2\alpha^2 + 1}{4\alpha^2} \); and then

\[
\tilde{\phi} = Q \cosh R, \quad \tilde{\sigma} = Q \sinh R.
\]

(4.132)

A conformal transformation \( \tilde{g}_{\mu\nu} = Q^2 g_{\mu\nu} \) leads to the Einstein frame action:

\[
S_{\text{EF}} = \frac{1}{2k\kappa_5^2(2\alpha^2 + 1)} \int d^4x \sqrt{-g} \left[ R - \frac{12\alpha^2}{1 + 2\alpha^2} \frac{(\partial Q)^2}{Q^2} \right. \\
- \left. \frac{6}{2\alpha^2 + 1} (\partial R)^2 \right].
\]

Note that in this frame both fields have the correct sign in front of the kinetic terms. For \( \alpha \to 0 \) (i.e., the RS case) the \( Q \)-field decouples. This reflects the fact, that the bulk scalar field decouples, and the only scalar degree of freedom is the distance between the branes. One can read off the gravitational constant to be

\[
16\pi G = \frac{2k\kappa_5^2}{1 + 2\alpha^2}.
\]

The matter sector of the action can be found easily: if matter lives on the branes, it 'feels' the induced metric. That is, the action has the form

\[
S_m^{(1)} = S_m^{(1)}(\Psi_1, g_{\mu\nu}^{B(1)}) \quad \text{and} \quad S_m^{(2)} = S_m^{(2)}(\Psi_2, g_{\mu\nu}^{B(2)}),
\]
where $g^{B(i)}_{\mu\nu}$ denotes the induced metric on each branes. In going to the Einstein frame one gets

$$S^{(1)}_m = S^{(1)}_m (\Psi_1, A^2(Q, R)g_{\mu\nu}) \quad \text{and} \quad S^{(2)}_m = S^{(2)}_m (\Psi_2, B^2(Q, R)g_{\mu\nu}),$$

where matter now couples explicitly to the fields via the functions $A$ and $B$, which we will give below (neglecting derivative interactions).

The theory derived with the help of the moduli space approximation has the form of a multi–scalar–tensor theory, in which matter on both branes couple differently to the moduli fields. We note, that methods different from the moduli–space approximation have been used in the literature in order to get the low–energy effective action or the resulting field equations for a two–brane system. Qualitatively, the features of the resulting theories agree with the moduli–space approximation discussed above.

In the following we will discuss observational constraints imposed on the parameter of the theory.

### Observational Constraints

In order to constrain the theory, it is convenient to write the moduli Lagrangian in the form of a non-linear sigma model with kinetic terms

$$\gamma_{ij} \partial \phi^i \partial \phi^j,$$

where $i = 1, 2$ labels the moduli $\phi^1 = Q$ and $\phi^2 = R$. The sigma model couplings are here

$$\gamma_{QQ} = \frac{12\alpha^2}{1 + 2\alpha^2} \frac{1}{Q^2}, \quad \gamma_{RR} = \frac{6}{1 + 2\alpha^2}.$$

Notice the potential danger of the $\alpha \to 0$ limit, the RS model, where the coupling to $Q$ becomes very small. In an ordinary Brans–Dicke theory with a single field, this would correspond to a vanishing Brans-Dicke parameter which is ruled out experimentally. Here we will see that the coupling to matter is such that this is not the case. Indeed we can write the action expressing the coupling to ordinary matter on our brane as

$$A = a(\phi) f^{-1/2}(\phi, \sigma), \quad B = a(\sigma) f^{-1/2}(\phi, \sigma),$$

where we have neglected the derivative interaction.

Let us introduce the parameters

$$\alpha_Q = \partial_Q \ln A, \quad \alpha_R = \partial_R \ln A.$$
We find that \( \lambda = \frac{4}{1 + 2\alpha^2} \) \[
A = Q - \frac{\alpha^2}{2} (\cosh R)^{\frac{1}{2}},
\]
leading to \[
\alpha_Q = -\frac{\alpha^2 \lambda}{2} \frac{1}{Q}, \quad \alpha_R = \frac{\lambda \tanh R}{4}.
\]
Observations constrain the parameter \( \theta = \gamma^{ij} \alpha_i \alpha_j \) to be less than \( 10^{-3} \) [Brax and Bruck (2003)]. We get therefore a bound on \[
\theta = \frac{4}{3} \frac{\alpha^2}{1 + 2\alpha^2} + \frac{\tanh^2 R}{6(1 + 2\alpha^2)}.
\]
The bound implies that \( \alpha \leq 10^{-2}, \quad R \leq 0.2. \)
The smallness of \( \alpha \) indicates a strongly warped bulk geometry such as an Anti–de Sitter space–time. In the case \( \alpha = 0 \), we can easily interpret the bound on \( R \). Indeed in that case \[
\tanh R = e^{-k(\sigma - \phi)},
\]
i.e., this is nothing but the exponential of the radion field measuring the distance between the branes. We get that gravity experiments require the branes to be sufficiently far apart. When \( \alpha \neq 0 \) but small, one way of obtaining a small value of \( R \) is for the hidden brane to become close from the would-be singularity where \( a(\sigma) = 0. \)

We would like to mention that the parameter \( \theta \) can be calculated also for matter on the negative tension brane. Then, following the same calculations as above, it can be seen that the observational constraint for \( \theta \) cannot be satisfied. Thus, if the standard model particles are confined on the negative tension brane, the moduli have necessarily to be stabilized. In the following we will assume that the standard model particles are confined on the positive tension brane and study the cosmological evolution of the moduli fields.
Cosmological Implications

The discussion in the last subsection raises an important question: the parameter $\alpha$ has to be chosen rather small, in order for the theory to be consistent with observations. Similarly the field $R$ has to be small too. The field $R$ is dynamical and one would like to know if the cosmological evolution drives the field $R$ to small values such that it is consistent with the observations today. Otherwise are there natural initial conditions for the field $R$? In the following we study the cosmological evolution of the system in order to answer these questions.

The field equations for a homogenous and isotropic universe can be obtained from the action. The Friedmann equation reads

$$H^2 = \frac{8\pi G}{3} (\rho_1 + \rho_2 + V_{\text{eff}} + W_{\text{eff}}) + \frac{2\alpha^2}{1 + 2\alpha^2} \dot{\phi}^2 + \frac{1}{1 + 2\alpha^2} R^2. \quad (4.133)$$

where we have defined $Q = \exp\phi$. The field equations for $R$ and $\phi$ read

$$\ddot{R} + 3H \dot{R} = -8\pi G \frac{1 + 2\alpha^2}{6} \left[ \frac{\partial V_{\text{eff}}}{\partial R} + \frac{\partial W_{\text{eff}}}{\partial R} \right] + \alpha^{(1)}_R (\rho_1 - 3p_1) + \alpha^{(2)}_R (\rho_2 - 3p_2) \right] \quad (4.134)$$

$$\ddot{\phi} + 3H \dot{\phi} = -8\pi G \frac{1 + 2\alpha^2}{12\alpha^2} \left[ \frac{\partial V_{\text{eff}}}{\partial \phi} + \frac{\partial W_{\text{eff}}}{\partial \phi} \right] + \alpha^{(1)}_\phi (\rho_1 - 3p_1) + \alpha^{(2)}_\phi (\rho_2 - 3p_2) \right]. \quad (4.135)$$

The coupling parameter are given by

$$\alpha^{(1)}_\phi = -\frac{2\alpha^2}{1 + 2\alpha^2}, \quad \alpha^{(2)}_\phi = -\frac{2\alpha^2}{1 + 2\alpha^2}, \quad (4.136)$$

$$\alpha^{(1)}_R = \frac{\tanh R}{1 + 2\alpha^2}, \quad \alpha^{(2)}_R = \frac{(\tanh R)^{-1}}{1 + 2\alpha^2}. \quad (4.137)$$

We have included matter on both branes as well as potentials $V_{\text{eff}}$ and $W_{\text{eff}}$ on each branes. We now concentrate on the case where matter is only on our brane. In the radiation dominated epoch the trace of the energy–momentum tensor vanishes, so that $Q$ and $\phi$ quickly become constant. The scale factor scales like $a(t) \propto t^{1/2}$. 
In the matter-dominated era, the solution to these equations is given by

$$\rho_1 = \rho_e \left( \frac{a}{a_e} \right)^{-3 - 2\alpha^2/3}$$
$$a = a_e \left( \frac{t}{t_e} \right)^{2/3 - 4\alpha^2/27},$$

together with

$$\phi = \phi_e + \frac{1}{3} \ln \frac{a}{a_e},$$
$$R = R_0 \left( \frac{t}{t_e} \right)^{-1/3} + R_1 \left( \frac{t}{t_e} \right)^{-2/3},$$
as soon as $t \gg t_e$. Note that $R$ indeed decays. This implies that small values of $R$ compatible with gravitational experiments are favoured by the cosmological evolution. Note, however, that the size of $R$ in the early universe is constrained by nucleosynthesis as well as by the CMB anisotropies.

A large discrepancy between the values of $R$ during nucleosynthesis and now induces a variation of the particle masses, or equivalently Newton’s constant, which is excluded experimentally. One can show that by putting matter on the negative tension brane as well, the field $R$ evolves even faster to zero. This behavior is reminiscent of the attractor solution in scalar–tensor theories [Brax and Bruck (2003)].

In the 5D picture the fact that $R$ is driven to small values means that the negative tension brane is driven towards the bulk singularity. In fact, solving the equations numerically for more general cases suggest that $R$ can even by negative, which is, in the 5D description meaningless, as the negative tension brane would move through the bulk singularity. Thus, in order to make any further progress, one has to understand the bulk singularity better. Of course, one could simply assume that the negative tension brane is destroyed when it hits the singularity. A more interesting alternative would be if the brane is repelled instead. It was speculated that this could be described by some effective potential in the low-energy effective action [Brax and Bruck (2003)].

4.3.8.5 *Brane Collision*

We have seen that brane world models are plagued with a singularity problem: the negative tension brane might hit a bulk singularity. In that case our description of the physics on the brane requires techniques beyond the field theory approach that we have followed in this review. It is only within

\footnote{For $\alpha = 0$ the theory is equivalent to the Randall–Sundrum model. In this case the bulk singularity is shifted towards the Anti–de Sitter boundary.}
a unified theory encompassing general relativity and quantum mechanics that such questions might be addressed. String theory may be such a theory. The problem of the nature of the resolution of cosmological singularities in string theory is still a vastly uncharted territory. There is a second kind of singularity which arises when two branes collide. In such a case there is also a singularity in the low energy effective action as one of the extra dimensions shrink to zero size. It was speculated that brane collisions play an important role in cosmology, especially in order to understand the Big–Bang itself.

As soon as the space–time contains several branes and that these branes move with respect to each other, they might collide. A fascinating possibility, which has been actively explored by [Khoury et al. (2001c); Bucher (2002); Gen et al. (2002)], is that the Big–Bang is such a brane collision. Rather than entering into the details of these various models, let us point out here a simple and general analysis [Langlois et al. (2002)] of the collision of (parallel or concentric) branes separated by vacuum, i.e., branes separated by patches of AdS–Schwarzschild space–times (allowing for different Schwarzschild–type mass and cosmological constant in each region) with the 5D AdS–Schwarzschild metric (AdS for $\Lambda < 0$, which is the case we are interested in; for $\Lambda > 0$, this the Schwarzschild–de Sitter metric)

$$ds^2 = -f(R)\,dT^2 + \frac{dR^2}{f(R)} + R^2\,\gamma_{ij}\,dx^i\,dx^j,$$

where (4.138)

$$f(R) \equiv k - \frac{\Lambda}{6} \, R^2 - \frac{C}{R^2}$$

(where $C$ is an arbitrary integration constant). Although we are interested here by 3-branes embedded in a 5D space–time, this analysis is immediately applicable to the case of $n$-branes moving in a $(n + 2)$-dimensional space–time, with the analogous symmetries.

To analyze the collision, it is convenient to introduce an angle $\alpha$, which characterizes the motion of the brane with respect to the coordinate system (4.138), defined by

$$\alpha = \sinh^{-1}(\epsilon R/\sqrt{f}),$$

where $\epsilon = +1$ if $R$ decreases from ‘left’ to ‘right’, $\epsilon = -1$ otherwise. Considering a collision involving a total number of $N$ branes, both ingoing and outgoing, thus separated by $N$ space–time regions, one can label alternately branes and regions by integers, starting from the leftmost ingoing brane and going anticlockwise around the point of collision. The branes will thus be
denoted by odd integers, $2k - 1$ ($1 \leq k \leq N$), and the regions by even integers, $2k$ ($1 \leq k \leq N$). Let us introduce the angle $\alpha_{2k-1|2k}$ which characterizes the motion of the brane $B_{2k-1}$ with respect to the region $R_{2k}$, and which is defined by [Langlois (2002)]

$$\sinh \alpha_{2k-1|2k} = \frac{\epsilon_{2k} R_{2k-1}}{\sqrt{f_{2k}}}. \quad (4.139)$$

Conversely, the motion of the region $R_{2k}$ with respect to the brane by the Lorentz angle

$$\alpha_{2k|2k-1} = -\alpha_{2k-1|2k}.$$ 

It can be shown that the junction conditions for the branes can be written in the form

$$\tilde{\rho}_{2k-1} \equiv \pm \frac{\kappa^2}{3} F_{2k-1} R = \epsilon_{2k} \sqrt{f_{2k}} \exp (\pm \alpha_{2k-1|2k})$$

$$- \epsilon_{2k-2} \sqrt{f_{2k-2}} \exp (\mp \alpha_{2k-2|2k-1}), \quad (4.140)$$

with the plus sign for ingoing branes ($1 \leq k \leq N$), the minus sign for outgoing branes ($N_{n} + 1 \leq k \leq N$). An outgoing positive energy density brane thus has the same sign as an ingoing negative energy density brane.

The advantage of this formalism becomes obvious when one writes the geometrical consistency relation that expresses the matching of all branes and space–time regions around the collision point. In terms of the angles defined above, it reads simply

$$\sum_{i=1}^{2N} \alpha_{j|i+1} = 0. \quad (4.141)$$

Moreover, introducing the generalized angles

$$\alpha_{j|j'} = \sum_{i=j}^{j'-1} \alpha_{i|i+1}, \quad \text{if} \quad j < j', \quad \text{and} \quad \alpha_{j'|j} = -\alpha_{j|j'},$$

the sum rule for angles (4.141) combined with the junction conditions (4.140) leads to the laws of energy conservation and momentum conservation. The energy conservation law reads [Langlois (2002)]

$$\sum_{k=1}^{N} \tilde{\rho}_{2k-1} \gamma_{j|2k-1} = 0,$$
where $\gamma_{jj'} \equiv \cosh \alpha_{jj'}$ corresponds to the Lorentz factor between the brane/region $j$ and the brane/region $j'$ and can be obtained, if $j$ and $j'$ are not adjacent, by combining all intermediary Lorentz factors (this is simply using the velocity addition rule of special relativity). The index $j$ corresponds to the reference frame with respect to which the conservation rule is written. Similarly, the momentum conservation law in the $j$th reference frame can be expressed in the form

$$\sum_{k=1}^{N} \tilde{p}_{2k-1} \gamma_{2k-1j} \beta_{2k-1j} = 0,$$

with $\gamma_{jj'} \beta_{jj'} \equiv \sinh \alpha_{jj'}$. One thus gets, just from geometrical considerations, conservation laws relating the brane energies densities and velocities before and after the collision point. These results apply to any collision of branes in vacuum, with the appropriate symmetries of homogeneity and isotropy. An interesting development would be to extend the analysis to branes with small perturbations and investigate whether one can find scenarios which can produce quasi–scale invariant adiabatic spectra, as seems required by current observations.

On the other hand, in heteroric M-theory the regime where the distance between the branes becomes small corresponds to the regime where the string coupling constant becomes small and therefore a perturbative heterotic treatment may be available. In particular for adiabatic processes the resulting small instanton transition has been thoroughly studied. Here we would like to present an analysis of such a collision and of the possible outcome of such a collision. A natural and intuitive phenomenon which may occur during a collision is the existence of a cosmological bounce. Such objects are not available in 4d under mild assumptions, and therefore can be exhibited as a purely extra dimensional signature [Brax and Bruck (2003)].

We will describe a $n$D theory with a scalar field and gravity whose solutions present a cosmological singularity at $t = 0$. It turns out that this model is the low energy approximation of a purely $(n+1)$D model where the extra dimension is an interval with two boundary branes. The singularity corresponds to the brane collision. In the $(n+1)$D picture, one can extend the motion of the branes past each other, hence providing a continuation of the brane motion after the collision. The $(n+1)$D space–time is equivalent to an orbifold where the identification between space–time points is provided by a Lorentz boost. These spaces are the simplest possible space–times with a singularity. As with ordinary spatial orbifolds, one may try to
define string theory in such backgrounds and analyse the stringy resolution of the singularity. Unfortunately, these orbifolds are not stable in general relativity ruling them out as candidate stringy backgrounds. Let us now briefly outline some of the arguments.

We have already investigated the moduli space approximation for models with a bulk scalar field. Here we will consider that at low energy the moduli space consists of a single scalar field $\phi$ coupled to gravity

$$S = \int d^n x \sqrt{-g} \left( R - \frac{1}{2} (\partial \phi)^2 \right).$$

(4.142)

Cosmological solutions with

$$ds^2 = a^2(t)[-dt^2 + dx_i dx_j]$$

can be easily obtained

$$a = a(1) |t|^{1/(n-2)}, \quad \phi = \phi(1) + \epsilon \sqrt{\frac{2(n-1)}{n-2}} \ln |t|,$$

(4.143)

where $\epsilon = \pm 1$. There are two branches corresponding to $t < 0$ and $t > 0$ connected by a singularity at $t = 0$.

So what is the extra dimensional origin of such a model? One can uplift the previous system to $(n+1)$ dimensions by defining

$$\psi = e^{\gamma \phi} \quad \text{and} \quad \bar{g}_{\mu\nu} = \psi^{-4/(n-2)} g_{\mu\nu},$$

where $\gamma = \sqrt{(n-2)/8(n-1)}$. Consider now the purely gravitational $(n+1)$ dimensional theory with the metric

$$ds^2_{n+1} = \psi^4 dw^2 + \bar{g}_{\mu\nu} dx^\mu dx^\nu,$$

where $w \in [0,1]$. The two boundaries at $w = 0$ and $w = 1$ are boundary branes. The dimensional reduction on the interval $w \in [0,1]$, i.e., integrating over the extra dimension, yields the effective action (4.142) provided one restricts the two fields $\psi(x^\mu)$ and $\bar{g}_{\mu\nu}(x^\mu)$ to be dependent on $nD$ only.

Let us now consider the nature of $(n+1)$ D space–time obtained from the solutions (4.143). The $(n+1)$D metric becomes

$$ds^2_{n+1} = B^2 t^2 dw^2 + \eta_{\mu\nu} dx^\mu dx^\nu,$$

for a given $B$ depending on the integrations constants $\phi(1)$ and $a(1)$. The geometry of space–time is remarkably simple. It is a direct product $R^{n-1} \times$
$M$ where $M$ is the two dimensional compactified Milne space whose metric is

$$ds^2_M = -dt^2 + B^4 t^2 dw^2.$$  

Using the light cone coordinates

$$x^\pm = \pm t e^{\pm B^2 w},$$

the metric of Milne space reads

$$ds^2_M = dx_+ dx_-,$$

coinciding with the two dimensional Minkowski metric. There is one subtlety here, the original identification of the extra–dimensional interval is here transcribed in the fact that Milne space is modded out by the boost

$$x^\pm \rightarrow e^{\pm 2 B^2} x^\pm,$$

as we have identified the interval with $S^1/Z_2$ and the boundary branes are the fixed points of the $Z_2$ action as in the RS model.

The two boundary branes collide at $x^\pm = 0$, their trajectories are given by

$$x_0^\pm = \pm t, \quad x_1^\pm = \pm t e^{\pm B^2}.$$  

At the singularity one can hope that the branes go past each other and evolve henceforth. Unfortunately, Horowitz and Polchinski have shown that the structure of the orbifold space–time is unstable [Horowitz and Polchinski (2002)]. By considering a particle in this geometry, they showed that space–time collapses to a space–like singularity. Indeed one can focus on a particle and its $n$th image under the orbifold action. In terms of collision the impact parameter $b$ becomes constant as $n$ grows while the center of mass energy $\sqrt{s}$ grows like $\cosh n B^2$. As soon as $n$ is large enough,

$$G \sqrt{s} > b^{n-2},$$

the two particle approach each other within their Schwarzschild radii therefore forming a black hole through gravitational collapse. So the orbifold space–time does not make sense in general relativity, i.e., not defining a time–dependent background for string theory.

Hence, it seems that the most simple example of brane collision needs to be modified in order to provide a working example of singularity with a meaningful string theoretic resolution. It would be extremely relevant
if one could find examples of stable backgrounds of string theory where a cosmological singularity can be resolved using stringy arguments. A particularly promising avenue is provided by S–branes where a cosmological singularity is shielded by a horizon [Brax and Bruck (2003)]. Time will certainly tell which of these approaches could lead to a proper understanding of cosmological singularities and their resolutions, an issue highly relevant to brane cosmology both in the early universe and the recent past.

4.3.8.6 Open Questions

In this section we have reviewed different aspects of brane cosmology in a hopefully pedagogical manner reflecting our own biased point of view. Let us finally summarize some of the open questions:

• In the case of the single brane model by Randall & Sundrum, the homogeneous cosmological evolution is well understood. An unsolved issue in this model, however, is a complete understanding of the evolution of cosmological perturbations. The effects of the bulk gravitational field, encoded in the projected Weyl–tensor, on CMB physics and Large Scale Structures are not known. The problem is twofold: first, the bulk equation are partial, nonlinear differential equations and second, boundary conditions on the brane have to be imposed. The current formalism have not yet been used in order to tackle these problems (for perturbations in brane world theories, see [Riazuelo (2002)].

• Models with bulk scalar fields: Although we have presented some results on the cosmological evolution of a homogeneous brane, we assumed that the bulk scalar field does not strongly vary around the brane. Clearly, this needs to be investigated in some detail through a detailed investigation of the bulk equations, presumably with the help of numerical methods. Furthermore, for models with two branes, the cosmology has to be explored also in the high energy regime, in which the moduli–space approximation is not valid. Some exact cosmological solutions have been found in [Lukas et al. (2000)].

• Both the bulk scalar field as well as the inter–brane distance in two brane models could play an important role at least during some part of the cosmological history. Maybe one of the fields plays the role of dark energy. In that case, it is only natural that masses of particles vary, as well as other parameter, such as the fine structure constant $\alpha_{em}$ [Brax et al. (2001)]. Details of this interesting proposal have yet
to be worked out.

- The bulk singularity, which was thought to be shielded away with the help of a second brane, seems to play an important role in a cosmological setting. We have seen that the negative tension brane moves towards the bulk singularity and eventually hits it. Therefore, cosmology forces us to think about this singularity, even if it was shielded with a second brane. Cosmologically, the brane might be repelled, which might be described by a potential. Alternatively, one might take quantum corrections into account in form of a Gauss–Bonnett term in the bulk.

- Brane collisions provide a different singularity problem in brane cosmology. String theory has to make progress in order to understand this singularity as well. From the cosmological point of view, the question is, if a transition between the brane collision can provide a new cosmological era and how cosmological perturbations evolve before and after the bounce.

For more details, see Brax and Bruck (2003); Langlois (2002).

### 4.3.9 Hawking’s Brane New World

In this subsection, following Hawking et al. (2000), we present S. Hawking’s Brane New World.

As seen above, Randall and Sundrum have suggested Randall and Sundrum (1999b) that 4D gravity may be recovered in the presence of an infinite fifth dimension provided that we live on a domain wall embedded in anti–de Sitter space (AdS). Their linearized analysis showed that there is a massless bound state of the graviton associated with such a wall as well as a continuum of massive Kaluza–Klein modes.\(^{42}\)

Recall that Kaluza–Klein theory (KK) is a model that seeks to unify the two fundamental forces of gravitation and electromagnetism. The theory was first published in 1921 and was discovered by the mathematician T. Kaluza who extended general relativity to a 5D space–time. The resulting equations can be separated out into further sets of equations, one of which is equivalent to Einstein field equations, another set equivalent to Maxwell’s equations for the electromagnetic field and the final part an extra scalar field now termed the radion.

In modern geometry, the extra fifth dimension can be understood to be the circle group $U(1)$, as electromagnetism can essentially be formulated as a gauge theory on a fiber bundle, the circle bundle, with gauge group $U(1)$. Once this geometrical interpretation is understood, it is relatively straightforward to replace $U(1)$ by a general Lie group, to get Yang–Mills theories. If a distinction is drawn, then it is that Yang–Mills theories occur on a flat space–time, whereas KK treats the more general case of curved space–time.
The base space of KK need not be 4D space–time; it can be any (pseudo)Riemannian manifold, or even a supersymmetric manifold or orbifold or even a noncommutative space.

As an approach to the unification of the forces, it is straightforward to apply the KK in an attempt to unify gravity with the strong and electroweak forces by using the symmetry group of the Standard Model, \( SU(3) \times SU(2) \times U(1) \). However, a naive attempt to convert this interesting geometrical construction into a bona-fide model of reality founders on a number of issues, including the fact that the fermions must be introduced in an artificial way (in nonsupersymmetric models). A less problematic approach to the unification of the forces is taken by modern string theory and M–theory. Nonetheless, KK remains an important touchstone in theoretical physics and is often embedded in more sophisticated theories. It is studied in its own right as an object of geometric interest in K–theory.

Even in the absence of a completely satisfying theoretical physics framework, the idea of exploring extra, compactified, dimensions is of considerable interest in the experimental physics and astrophysics communities. A variety of predictions, with real experimental consequences, can be made (in the case of large extra dimensions/warped models). For example, on the simplest of principles, one might expect to have standing waves in the extra compactified dimension(s). If an extra dimension is of radius \( R \), the energy of such a standing wave would (naively) be \( E = nhc/R \) with \( n \) an integer, \( h \) being Planck’s constant and \( c \) the speed of light. This set of possible energy values is often called the Kaluza–Klein tower.

To build the Kaluza–Klein theory, one picks an invariant metric on the circle \( S^1 \) that is the fiber of the \( U(1) \)–bundle of electromagnetism. Suppose this metric gives the circle a total length of \( \Lambda \). One then considers metrics \( \hat{g} \) on the bundle \( P \) that are consistent with both the fiber metric, and the metric on the underlying manifold \( M \). The consistency conditions are:

(i) the projection of \( \hat{g} \) to the vertical subspace \( \text{Vert}_p P \subset T_p P \) needs to agree with metric on the fiber over a point in the manifold \( M \), and

(ii) the projection of \( \hat{g} \) to the horizontal subspace \( \text{Hor}_p P \subset T_p P \) of the tangent space at point \( p \in P \) must be isomorphic to the metric \( g \) on \( M \) at \( \pi(p) \).

The \textit{Kaluza–Klein action} for such a metric is given by

\[ S(\hat{g}) = \int_P R(\hat{g}) \, \text{vol}(\hat{g}). \]

The scalar curvature, written in components, then expands to

\[ R(\hat{g}) = \pi^* \left( R(g) - \frac{\Lambda^2}{2} |F|^2 \right), \]

where \( \pi^* \) is the pull–back of the fiber–bundle projection \( \pi : P \to M \). The connection \( A \) on the fiber bundle is related to the electromagnetic field strength as

\[ \pi^* F = dA. \]

That there always exists such a connection, even for fiber bundles of arbitrarily complex topology, is a result from homology and specifically, \textit{K–theory}. Applying \textit{Fubini’s Theorem} and integrating on the fiber, one gets

\[ S(\hat{g}) = \Lambda \int_M \left( R(g) - \frac{1}{\Lambda^2} |F|^2 \right) \, \text{vol}(g). \]
RS used *horospherical coordinates* based on slicing AdS into flat hypersurfaces. An issue that has not received much attention so far is the role of boundary conditions at these Cauchy horizons in AdS. With stationary perturbations, one can impose the boundary conditions that the horizons remain regular. Indeed, without this boundary condition the solution for stationary perturbations is not well defined. Even for non-perturbative departures from the RS solution, like black holes, one can impose the boundary condition that the AdS horizons remain regular [Chamblin *et al.* (2000); *Emparan et al.* (2000)]. Non–stationary perturbations on the domain wall, however, will give rise to gravitational waves that cross the horizons. This will tend to focus the null geodesic generators of the horizon, which will mean that they will intersect each other on some caustic. Beyond the caustic, the null geodesics will not lie in the horizon. However, null geodesic generators of the future event horizon cannot have a future endpoint [Hawking and Ellis (1973)] and so the endpoint must lie to the past. We conclude that if the past and future horizons remain non–singular when perturbed (as required for a well–defined boundary condition) then they must intersect at a finite distance from the wall. By contrast, the past and future horizons don’t intersect in the RS ground state but go off to infinity in AdS.

The RS horizons are like the horizons of extreme black holes. When considering perturbations of black holes, one normally assumes that radiation

Varying the action with respect to the component $A$, one regains the Maxwell equations. Applying the variational principle to the base metric $g$, one gets the *Einstein equation*

$$R_{ij} - \frac{1}{2} g_{ij} R = \frac{1}{\Lambda^2} T_{ij},$$

with the *Maxwell stress–energy tensor* being given by

$$T^{ij} = F^{(i} g^{j)} - \frac{1}{4} g^{ij} |F|^2.$$

The original theory identifies $\Lambda$ with the fiber metric $g_{55}$, and allows $\Lambda$ to vary from fiber to fiber. In this case, the coupling between gravity and the electromagnetic field is not constant, but has its own dynamical field, the radion.

In the above, the *size* of the loop $\Lambda$ acts as a coupling constant between the gravitational field and the electromagnetic field. If the base manifold is 4D, the *Kaluza–Klein manifold* $P$ is 5D. The fifth dimension is a *compact space*, and is called the compact dimension. The phenomenon of having a higher-dimensional manifold where some of the dimensions are compact is referred to as *compactification*.

The above development generalizes in a more-or-less straightforward fashion to general principal $G$–bundles for some arbitrary Lie group $G$ taking the place of $U(1)$. In such a case, the theory is often referred to as a *Yang–Mills theory*, and is sometimes taken to be synonymous. If the underlying manifold is supersymmetric, the resulting theory is a *supersymmetric Yang–Mills theory*. 
Quantum Leap

can flow across the future horizon but that nothing comes out of the past horizon. This is because the past horizon is not really there, and should be replaced by the collapse that formed the black hole. To justify a similar boundary condition on the Randall-Sundrum past horizon, one needs to consider the initial conditions of the universe.

The main contender for a theory of initial conditions is the no boundary proposal [Hartle and Hawking (1983)] that the quantum state of the universe is given by a Euclidean path integral over compact metrics. The simplest way to implement this proposal for the Randall Sundrum idea is to take the Euclidean version of the wall to be a four sphere at which two balls of $AdS_5$ are joined together. In other words, take two balls in $AdS_5$, and glue them together along their four sphere boundaries. The result is topologically a five sphere, with a delta function of curvature on a 4D domain wall separating the two hemispheres. If one analytically continues to Lorentzian signature, one gets a 4D de Sitter hyperboloid, embedded in Lorentzian anti de Sitter space.

The past and future RS horizons, are replaced by the past and future light cones of the points at the centers of the two balls. Note that the past and future horizons now intersect each other and are non extreme, which means they are stable to small perturbations. A perfectly spherical Euclidean domain wall will give rise to a 4D Lorentzian universe that expands forever in an inflationary manner.\footnote{Such inflationary brane-world solutions have been studied in [Chamblin and Reall (1999); Kaloper (1999); Nihei (1999)].}

In order for a spherical domain wall solution to exist, the tension of the wall must be larger than the value assumed by RS, who had a flat domain wall. We shall assume that matter on the wall increases its effective tension, permitting a spherical solution. Below we consider a strongly coupled large $N$ CFT on the domain wall. On a spherical domain wall, the conformal anomaly of the CFT increases the effective tension of the domain wall, making the spherical solution possible. The Lorentzian geometry is a de Sitter universe with the conformal anomaly driving inflation.

The no boundary proposal allows one to calculate unambiguously the graviton correlator on the domain wall. In particular, the Euclidean path integral itself uniquely specifies the allowed fluctuation modes, because perturbations that have infinite Euclidean action are suppressed in the path integral. Therefore, in this framework, there is no need to impose by hand an additional, external prescription for the vacuum state for each perturbation mode. In addition, the AdS/CFT correspondence allows a fully quan-
tum mechanical treatment of the CFT, in contrast with the usual classical
treatment of matter fields in inflationary cosmology.

Finally, we analytically continue the Euclidean correlator into the
Lorentzian region, where it describes the spectrum of quantum mechanical
vacuum fluctuations of the graviton field on an inflating domain wall
with conformally invariant matter living on it. We find that the quantum
loops of the large $N$ CFT give space–time a rigidity that strongly suppresses
metric fluctuations on small scales. Since any matter would be expected to
behave like a CFT at small scales, this result probably extends to any infla-
tionary model with sufficiently many matter fields. It has long been known
that matter loops lead to short distance modifications of gravity. Our work
shows that these modifications can lead to observable consequences in an
inflationary scenario.

Although we have carried out our calculations for the RS model, we
shall show that results for 4D Einstein gravity coupled to the CFT can be
recovered by taking the domain wall to be large compared with the AdS
scale. Thus our conclusion that metric fluctuations are suppressed holds
independently of the RS scenario.

The spherical domain wall considered in this section analytically con-
tinues to a Lorentzian de Sitter universe that inflates forever. How-
ever, Starobinsky [Starobinsky (1980)] showed that the conformal anomaly
driven de Sitter phase is unstable to evolution into a matter dominated uni-
verse. If such a solution could be obtained from a Euclidean instanton then
it would have an $O(4)$–symmetry group, rather than the $O(5)$–symmetry
of a spherical instanton.

The AdS/CFT correspondence [Maldacena (1998); Gubser et al. (1998);
Witten (1998b)] provides an explanation of the RS behavior. It relates the
RS model to an equivalent 4D theory consisting of general relativity coupled
to a strongly interacting conformal field theory and a logarithmic correction.
Under certain circumstances, the effects of the CFT and logarithmic term
are negligible and pure gravity is recovered.

4.3.9.1 RS Scenario from AdS/CFT

The AdS/CFT correspondence [Maldacena (1998); Gubser et al. (1998);
Witten (1998b)] relates IIB supergravity theory in $AdS_5 \times S^5$ to a $\mathcal{N} = 4$
$U(N)$ superconformal field theory. If $g_{YM}$ is the coupling constant of this
theory then the ’t Hooft parameter is defined to be $\lambda = g_{YM}^2 N$. The

\footnote{This was first pointed out in unpublished remarks of Maldacena and Witten.}
CFT parameters are related to the supergravity parameters by Maldacena (1998)

\[ l = \lambda^{1/4} l_s, \quad \frac{l^3}{G} = \frac{2N^2}{\pi}, \tag{4.144} \]

where \( l_s \) is the string length, \( l \) the AdS radius and \( G \) the 5D Newton constant. Note that \( \lambda \) and \( N \) must be large in order for stringy effects to be small. The CFT lives on the conformal boundary of \( \text{AdS}_5 \). The correspondence takes the following form [Hawking et al. (2000)]:

\[ Z[h] \equiv \int \mathcal{D}[g] \exp(-S_{\text{grav}}[g]) = \int \mathcal{D}[\phi] \exp(-S_{\text{CFT}}[\phi,h]) \equiv \exp(-W_{\text{CFT}}[h]), \tag{4.145} \]

here \( Z[h] \) denotes the supergravity partition function in \( \text{AdS}_5 \). This is given by a path integral over all metrics in \( \text{AdS}_5 \) which induce a given conformal equivalence class of metrics \( h \) on the conformal boundary of \( \text{AdS}_5 \). The correspondence relates this to the generating functional \( W_{\text{CFT}} \) of connected Green’s functions for the CFT on this boundary. This functional is given by a path integral over the fields of the CFT, denoted schematically by \( \phi \). Other fields of the supergravity theory can be included on the left hand side; these act as sources for operators of the CFT on the right hand side.

A problem with equation (4.145) as it stands is that the usual gravitational action in AdS is divergent, rendering the path integral ill-defined. A procedure for solving this problem was developed in [Witten (1998b); Tseytlin and Liu (1998); Henningsson and Skenderis (1998); Balasubramanian and Kraus (1999); Emparan et al. (1999); Kraus et al. (1999)]. First one brings the boundary into a finite radius. Next one adds a finite number of counterterms to the action in order to render it finite as the boundary is moved back off to infinity. These counterterms can be expressed solely in terms of the geometry of the boundary. The total gravitational action for \( \text{AdS}_{n+1} \) becomes

\[ S_{\text{grav}} = S_{\text{EH}} + S_{\text{GH}} + S_1 + S_2 + \ldots \]

The first term is the usual Einstein–Hilbert action with a negative cosmological constant:

\[ S_{\text{EH}} = -\frac{1}{16\pi G} \int d^{n+1}x \sqrt{g} \left( R + \frac{n(n-1)}{l^2} \right) \]

\[ \text{45We use a positive signature metric and a curvature convention for which a sphere has positive Ricci scalar.} \]
the overall minus sign arises because we are considering a Euclidean theory. The second term in the action is the Gibbons–Hawking boundary term, which is necessary for a well-defined variational problem [Gibbons and Hawking (1977)]:

\[ S_{GH} = -\frac{1}{8\pi G} \int d^n x \sqrt{h} K, \]

where \( K \) is the trace of the extrinsic curvature of the boundary and \( h \) the determinant of the induced metric. The first two counterterms are given by the following [Balasubramanian and Kraus (1999); Emparan et al. (1999); Kraus et al. (1999)] (we use the results of Kraus et al. (1999) rotated to Euclidean signature)

\[ S_1 = \frac{n-1}{8\pi G} \int d^n x \sqrt{h}, \quad \text{and} \quad S_2 = \frac{l}{16\pi G(n-2)} \int d^n x \sqrt{h} R, \]

where \( R \) now refers to the Ricci scalar of the boundary metric. The third counterterm is

\[ S_3 = \frac{l^3}{16\pi G(n-2)^2(n-4)} \int d^n x \sqrt{h} \left( R_{ij} R^{ij} - \frac{n}{4(n-1)} R^2 \right), \quad (4.146) \]

where \( R_{ij} \) is the Ricci tensor of the boundary metric and boundary indices \( i, j \) are raised and lowered with the boundary metric \( h_{ij} \). This expression is ill-defined for \( n = 4 \), which is the case of most interest to us. With just the first two counter-terms, the gravitational action exhibits logarithmic divergences [Tseytlin and Liu (1998); Henningson and Skenderis (1998)] so a third term is needed. This term cannot be written solely in terms of a polynomial in scalar invariants of the induced metric and curvature tensors; it makes explicit reference to the cut-off (i.e., the finite radius to which the boundary is brought before taking the limit in which it tends to infinity). The form of this term is the same as (4.146) with the divergent factor of \( 1/(n-4) \) replaced by \( \log(R/\rho) \), where \( R \) measure the boundary radius and \( \rho \) is some finite renormalization length scale.

Following [Gubser et al. (1998)], we can now use the AdS/CFT correspondence to explain the behavior discovered by Randall and Sundrum. The (Euclidean) RS model has the following action:

\[ S_{RS} = S_{EH} + S_{GH} + 2S_1 + S_3. \]

---

46Our convention is the following. Let \( n \) denotes the outward unit normal to the boundary. The extrinsic curvature is defined as \( K_{\mu\nu} = h^\rho_{[\mu} b^\nu_{\rho]} n_{\sigma}, \) where \( b^\mu_{\nu} = \delta^\mu_{\nu} - n_{\nu} n^\mu \) projects quantities onto the boundary.
Here $2S_1$ is the action of a domain wall with tension $(n-1)/(4\pi Gl)$. The final term is the action for any matter present on the domain wall. The domain wall tension can cancel the effect of the bulk cosmological constant to produce a flat domain wall. However, we are interested in a spherical domain wall so we assume that the matter on the wall gives an extra contribution to the effective tension. We shall discuss a specific candidate for the matter on the wall later on. The wall separates two balls $B_1$ and $B_2$ of $AdS$.

We want to study quantum fluctuations of the metric on the domain wall. Let $g_0$ denote the 5D background metric we have just described and $h_0$ the metric it induces on the wall. Let $h$ denote a metric perturbation on the wall. If we wish to calculate correlators of $h$ on the domain wall then we are interested in a path integral of the form \[ \langle h_{ij}(x)h_{i'j'}(x') \rangle = \int D[h] Z[h] h_{ij}(x)h_{i'j'}(x'), \] where

\[
Z[h] = \int_{B_1 \cup B_2} D[\delta g] D[\phi] \exp(-S_{RS}[g_0 + \delta g]) \times \int_{B_1 \cup B_2} D[\delta g] D[\phi] \exp(-S_{EH}[g_0 + \delta g] - S_{GH}[g_0 + \delta g] - S_m[\phi; h_0 + h]),
\]

$\delta g$ denotes a metric perturbation in the bulk that approaches $h$ on the boundary and $\phi$ denotes the matter fields on the domain wall. The integrals in the two balls are independent so we can replace the path integral by

\[
Z[h] = \exp(-2S_1[h_0 + h]) \times \left( \int_{B} D[\delta g] \exp(-S_{EH}[g_0 + \delta g] - S_{GH}[g_0 + \delta g]) \right)^2 \times \int_{B} D[\phi] \exp(-S_m[\phi; h_0 + h]),
\]

where $B$ denotes either ball. We now take $n = 4$ and use the AdS/CFT correspondence \[47\] to replace the path integral over $\delta g$ by the generating

---

[47] In principle, we should worry about gauge fixing and ghost contributions to the gravitational action. A convenient gauge to use in the bulk is transverse traceless gauge. We shall only deal with metric perturbations that also appear transverse and traceless on the domain wall. The gauge fixing terms vanish for such perturbations and the ghosts only couple to these perturbations at higher orders.
functional for a conformal field theory:

\[
\int_B \mathcal{D}[\delta g] \exp(-S_{EH}[g_0 + \delta g] - S_{GH}[g_0 + \delta g]) = \exp(-W_{RS}[h_0 + h] + S_1[h_0 + h] + S_2[h_0 + h] + S_3[h_0 + h]),
\]

we shall refer to this CFT as the RS–CFT since it arises as the dual of the RS geometry. It has gauge group \(U(N_{RS})\), where \(N_{RS}\) is given by equation (4.144). Strictly speaking, we are using an extended form of the AdS/CFT conjecture, which asserts that supergravity theory in a finite region of AdS is dual to a CFT on the boundary of that region with an ultraviolet cut-off related to the radius of the boundary. The path integral for the metric perturbation becomes

\[
Z[h] = \exp(-2W_{RS}[h_0 + h] + 2S_2[h_0 + h] + 2S_3[h_0 + h]) \int \mathcal{D}[\phi] \exp(-S_m[\phi; h_0 + h]).
\]

The RS model has been replaced by a CFT and a coupling to matter fields and the domain wall metric given by the action

\[-2S_2[h_0 + h] - 2S_3[h_0 + h] + S_m[\phi; h_0 + h].\]

The remarkable feature of this expression is that the term \(-2S_2\) is precisely the (Euclidean) Einstein–Hilbert action for 4D gravity with a Newton constant given by the RS value \(G_4 = G/l\). Therefore the RS model is equivalent to 4D gravity coupled to a CFT with corrections to gravity coming from the third counter term. This explains why gravity is trapped to the domain wall.

At first sight this appears rather amazing. We started off with a quite complicated 5D system and have argued that it is dual to 4D Einstein gravity with some corrections and matter fields. However in order to use this description, we have to know how to calculate with the RS–CFT. At present, the only way we know of doing this is via AdS/CFT, i.e., going back to the 5D description. The point of the AdS/CFT argument is to explain why the RS ‘alternative to compactification’ works and also to explain the origin of the corrections to Einstein gravity in the RS model. Note that if the matter on the domain wall dominates the RS–CFT and the third counterterm then these can be neglected and a purely 4D description is adequate.
4.3.9.2 CFT on the Domain Wall

Long ago, Starobinsky studied the cosmology of a universe containing conformally coupled matter [Starobinsky (1980)]. CFTs generally exhibit a conformal anomaly when coupled to gravity (for a review, see Duff (1994)). Starobinsky gave a de Sitter solution in which the anomaly provides the cosmological constant. By analyzing homogeneous perturbations of this model, he showed that the de Sitter phase is unstable but could be long lived, eventually decaying to a FRW cosmology.

Here we will consider the RS analogue of Starobinsky’s model by putting a CFT on the domain wall. On a spherical domain wall, the conformal anomaly provides the extra tension required to satisfy the Israel equations. It is appealing to choose the new CFT to be a $N=4$ superconformal field theory because then the AdS/CFT correspondence makes calculations relatively easy. This requires that the CFT is strongly coupled, in contrast with Starobinsky’s analysis.

Our 5D Euclidean action is the following [Hawking et al. (2000)]:

$$S = S_{EH} + S_{GH} + 2S_1 + W_{CFT}. \quad (4.147)$$

We seek a solution in which two balls of $AdS_5$ are separated by a spherical domain wall. Inside each ball, the metric can be written

$$ds^2 = l^2(dy^2 + \sinh^2 y d\Omega_n^2), \quad \text{with} \quad 0 \leq y \leq y_0.$$

The domain wall is at $y = y_0$ and has radius: $R = l \sinh y_0$. The effective tension of the domain wall is given by the Israel equations as

$$\sigma_{eff} = \frac{3}{4\pi G l} \coth y_0.$$

The actual tension of the domain wall is $\sigma = 3/(4\pi G l)$. We therefore need a contribution to the effective tension from the CFT. This is provided by the conformal anomaly, which takes the value $\langle T \rangle = -\frac{3N^2}{8\pi^2 R^4}$.

48 We emphasize that this use of the AdS/CFT correspondence is independent of the use described above because this new CFT is unrelated to the RS CFT [Hawking et al. (2000)].

49 Note that the conformal anomaly is the same at strong and weak coupling [Henningson and Skenderis (1998)] so any differences arising from strong coupling can only show up when we perturb the system.
This contributes an effective tension $-\langle T \rangle / 4$. We can now get an equation for the radius of the domain wall:

$$\frac{R^3}{l^3} \sqrt{\frac{R^2}{l^2} + 1} = \frac{N^2 G}{8\pi l^8} + \frac{R^4}{l^4}. \tag{4.148}$$

It is easy to see that this has a unique positive solution for $R$.

We are particularly interested in how perturbations of this model would appear to inhabitants of the domain wall. Thus we are interested in metric perturbations on the sphere

$$ds^2 = (R^2 \hat{\gamma}_{ij} + h_{ij}) dx^i dx^j.$$

Here $\hat{\gamma}_{ij}$ is the metric on a unit $n-$sphere. We shall only consider tensor perturbations, for which $h_{ij}$ is transverse and traceless with respect to $\hat{\gamma}_{ij}$.

In order to calculate correlators of the metric perturbation, we need to know the action to second order in the perturbation. The most difficult part here is obtaining $W_{CFT}$ to second order.

### CFT Generating Function

We want to work out the effect of the perturbation on the CFT on the sphere. To do this we use AdS/CFT. Introduce a fictional AdS region that fills in the sphere. Let $\bar{l}, \bar{G}$ be the AdS radius and Newton constant of this region. We emphasize that this region has nothing to do with the regions of AdS that ‘really’ lie inside the sphere in the RS scenario. This new AdS region is bounded by the sphere. If we take $\bar{l}$ to zero then the sphere is effectively at infinity in AdS so we can use AdS/CFT to calculate the generating functional of the CFT on the sphere. In other words, $\bar{l}$ is acting like a cut–off in the CFT and taking it to zero corresponds to removing the cut–off. However the relation

$$\frac{\bar{l}^3}{\bar{G}} = \frac{2N^2}{\pi}, \tag{4.149}$$

implies that if $\bar{l}$ is taken to zero then we must also take $\bar{G}$ to zero since $N$ is fixed (and large).

For the unperturbed sphere, the metric in the new AdS region is

$$ds^2 = \bar{l}^2 (dy^2 + \sinh^2 y \hat{\gamma}_{ij} dx^i dx^j),$$

and the sphere is at $y = y_0$ given by $R = \bar{l} \sinh y_0$. Note that $y_0 \to \infty$ as $\bar{l} \to 0$ since $R$ is fixed. In order to use AdS/CFT for the perturbed sphere, we need to know how the perturbation extends into the bulk. This is done...
by solving the linearized Einstein equations. It is always possible to choose a gauge in which the bulk metric perturbation takes the form
\[ h_{ij}(y, x) \, dx^i \, dx^j, \]
where \( h_{ij} \) is transverse and traceless with respect to the metric on the spherical spatial sections:
\[ \hat{\gamma}^{ij}(x) h_{ij}(y, x) = \hat{\nabla}^i h_{ij}(y, x) = 0, \]
with \( \hat{\nabla} \) denoting the covariant derivative defined by the metric \( \hat{\gamma}^{ij} \). Since we are only dealing with tensor perturbations, this choice of gauge is consistent with the boundary sitting at constant \( y \). If scalar metric perturbations were included then we would have to take account of a perturbation in the position of the boundary.

The linearized Einstein equations in the bulk are (for any dimension) [Hawking et al. (2000)]
\[ \nabla^2 h_{\mu\nu} = -\frac{2}{l^2} h_{\mu\nu}, \]
(4.150)
where \( \mu, \nu \) are \( n + 1 \)D indices. It is convenient to expand the metric perturbation in terms of tensor spherical harmonics \( H^{(p)}_{ij}(x) \). These obey
\[ \hat{\nabla}^2 H^{(p)}_{ij}(x) = \hat{\nabla}^2 H^{(p)}_{ij}(x) = 0, \]
and they are tensor eigenfunctions of the Laplacian:
\[ \hat{\nabla}^2 H^{(p)}_{ij} = (2 - p(p + n - 1)) \, H^{(p)}_{ij}, \]
where \( p = 2, 3, \ldots \). We have suppressed extra labels \( k, l, m, \ldots \) on these harmonics. The harmonics are orthonormal with respect to the obvious inner product. See [Higuchi (1987)] for more details of their properties. The metric perturbation can be written as a sum of separable perturbations of the form
\[ h_{ij}(y, x) = f_p(y) H^{(p)}_{ij}(x). \]
Substituting this into equation (4.150) gives
\[ f''_p(y) + (n - 4) \coth y f'_p(y) - (2(n - 2) + p(p + n - 1) + 2(n - 3)\cosech^2 y) f_p(y) = 0. \]
(4.151)
The roots of the indicial equation are $p + 2$ and $-p - n + 3$, yielding two linearly independent solutions for each $p$. In order to compute the generating functional $W_{\text{CFT}}$ we have to calculate the Euclidean action of these solutions. However, because the latter solution goes as $y^{-(p+n-3)}$ at the origin $y = 0$ of the instanton, the corresponding fluctuation modes have infinite Euclidean action\footnote{This can be seen by surrounding the origin by a small sphere $y = \epsilon$ and calculating the surface terms in the action that arise on this sphere. They are the same as the surface terms in equations (4.155) and (4.156) below, which are obviously divergent for the modes in question.}. Hence they are suppressed in the path integral. Therefore, in contrast to the methods where one requires a (rather ad hoc) prescription for the vacuum state of each perturbation mode, there is no need to impose boundary conditions by hand in our approach: the Euclidean path integral defines its own boundary conditions, which automatically gives a unique Green function. The path integral unambiguously specifies the allowed fluctuation modes as those which vanish at $y = 0$. Note that boundary conditions at the origin in Euclidean space replace the need for boundary conditions at the horizon in Lorentzian space. The solution regular at $y = 0$ is given by

$$f_p(y) = \frac{\sinh^{p+2}y}{\cosh^p y} F(p/2, (p + 1)/2, p + (n + 1)/2, \tanh^2 y).$$

This solution can also be written in terms of associated Legendre functions:

$$f_p(y) \propto (\sinh y)^{(5-n)/2} P_{-(n+1)/2}^{(p+n-1)/2} (\cosh y)$$

$$\propto (\sinh y)^{(4-n)/2} Q_{p+(n-2)/2}^{n/2} (\coth y),$$

and the latter can be related to Legendre functions if $n/2$ is an integer, using

$$Q^m_n(z) = (z^2 - 1)^{m/2} n^m Q_m \frac{d^m z}{dz}. $$

The full solution for the metric perturbation is

$$h_{ij}(y, x) = \sum_p \frac{f_p(y)}{f_p(y_0)} H^{(p)}_{ij}(x) \int d^n x' \sqrt{g} h_{kl}(x') H^{(p)}_{kl}(x').$$

We have a solution for the metric perturbation throughout the bulk region. The AdS/CFT correspondence can now be used to give the generating functional of the CFT on the perturbed sphere:

$$W_{\text{CFT}} = S_{\text{EH}} + S_{\text{GH}} + S_1 + S_2 + \ldots.$$
We shall give the terms on the right hand side for \( n = 4 \).

The Einstein–Hilbert action with cosmological constant is

\[
S_{EH} = -\frac{1}{16\pi G} \int d^5x \sqrt{\bar{g}} \left( R + \frac{12}{l^2} \right),
\]

and perturbing this gives

\[
S_{bulk} = -\frac{1}{16\pi G} \int d^5x \sqrt{\bar{g}} \left( -\frac{8}{l^2} + \frac{1}{4} h^{\mu\nu} \nabla_\mu h_{\nu}\rho + \frac{1}{2l^2} h^{\mu\nu} h_{\mu\nu} \right)
- \frac{1}{16\pi G} \int d^4x \sqrt{\gamma} \left( -\frac{1}{2} n^{\mu\nu\rho} \nabla_\nu h_{\mu}\rho + \frac{3}{4} h_{\mu\rho} n^{\mu\nu} \nabla_\nu h^{\rho\nu} \right),
\]

where Greek indices are 5D and we are raising and lowering with the unperturbed 5D metric. \( n = l dy \) is the unit normal to the boundary and \( \nabla \) is the covariant derivative defined with the unperturbed bulk metric. \( \gamma_{ij} = R^2 \bar{\gamma}_{ij} \) is the unperturbed boundary metric. It is important to keep track of all the boundary terms arising from integration by parts. Evaluating on shell gives [Hawking et al. (2000)]

\[
S_{EH} = \frac{\bar{l}^3}{2\pi G} \int d^4x \sqrt{\bar{\gamma}} \int_0^{\gamma_0} dy \sinh^4 y \quad (4.155)
- \frac{\bar{l}^3}{16\pi G} \int d^4x \sqrt{\bar{\gamma}} \left( \frac{3}{4l^2} h^{ij} \partial_y h_{ij} - \frac{\coth y_0}{l^4} h^{ij} h_{ij} \right)
\]

where we are now raising and lowering with \( \bar{\gamma}_{ij} \). The Gibbons–Hawking term is

\[
S_{GH} = -\frac{\bar{l}^3}{2\pi G} \int d^4x \sqrt{\bar{\gamma}} \left( \sinh^4 y_0 \cosh y_0 - \frac{1}{8l^2} h^{ij} \partial_y h_{ij} \right). \quad (4.156)
\]

The first counter term is

\[
S_1 = \frac{3}{8\pi G} \int d^4x \sqrt{\bar{\gamma}} \quad (4.157)
= \frac{3\bar{l}^3}{8\pi G} \int d^4x \sqrt{\bar{\gamma}} \left( \sinh^4 y_0 - \frac{1}{4l^2} h^{ij} h_{ij} \right).
\]

The second counter term is

\[
S_2 = \frac{\bar{l}}{32\pi G} \int d^4x \sqrt{\bar{\gamma}} R
= \frac{\bar{l}^3}{32\pi G} \int d^4x \sqrt{\bar{\gamma}} \left( 12 \sinh^2 y_0 - \frac{2}{l^2} \sinh^2 y_0 h^{ij} h_{ij} + \frac{1}{4l^4} \sinh^2 y_0 h^{ij} \nabla^2 h_{ij} \right).
\]
Thus with only two counter terms we would have

\[
W_{CFT} = \frac{3N^2 \Omega_4}{8\pi^2} \log \frac{R}{\bar{l}} 
- \frac{\bar{l}^4}{16\pi G} \int d^4x\sqrt{g} \left( -\frac{1}{4l^4} h^{ij} \partial_y h_{ij} + \frac{1}{l^4} h^{ij} h_{ij} \left( \frac{3}{2} - \sqrt{1 + \frac{l^2}{R^2}} \right) 
+ \frac{1}{l^4 R^2} f_{ij} h_{ij} - \frac{1}{8l^2 R^2} f^{ij} \nabla h_{ij} \right).
\]

\( \Omega_4 \) is the area of a unit 4–sphere and we have used equation (4.149). The expansion of \( \partial_y h_{ij} \) at \( y = y_0 \) is obtained from

\[
\partial_y h_{ij} = \sum_p f_p(y_0) H_{ij}^{(p)}(x) \int d^4x' \sqrt{g} h^{kl}(x') H_{kl}^{(p)}(x') \quad \text{and}
\]

\[
f_p(y_0) = 2 + \frac{\bar{l}^4}{2R^4} (p+1)(p+2) + p(p+1)(p+2)(p+3) \frac{\bar{l}^4}{4R^4} \log(\bar{l}/R) 
+ \frac{\bar{l}^4}{8R^4} [p^4 + 2p^3 - 5p^2 - 10p - 2 - p(p+1)(p+2)(p+3)(\psi(1) 
+ \psi(2) - \psi(p/2 + 2) - \psi(p/2 + 5/2))] + O \left( \frac{\bar{l}^6}{R^6} \log(\bar{l}/R) \right).
\]

The psi function is defined by \( \psi(z) = \Gamma'(z)/\Gamma(z) \). Substituting into the action we find that the divergences as \( \bar{l} \to 0 \) cancel at order \( R^4/\bar{l}^4 \) and \( R^2/\bar{l}^2 \).

The term of order \( \bar{l}^4/R^4 \) in the above expansion makes a contribution to the finite part of the action:

\[
W_{CFT} = \frac{3N^2 \Omega_4}{8\pi^2} \log \frac{R}{\bar{l}} + \frac{N^2}{256\pi^2 R^4} \sum_p \left( \int d^4x' \sqrt{g} h^{kl}(x') H_{kl}^{(p)}(x') \right)^2 \times (2p(p+1)(p+2)(p+3) \log(\bar{l}/R) + \psi(p) ,
\]

where

\[
\psi(p) = p(p+1)(p+2)(p+3)[\psi(p/2 + 5/2) + \psi(p/2 + 2) 
- \psi(2) - \psi(1)] + p^4 + 2p^3 - 5p^2 - 10p - 6.
\]

To cancel the logarithmic divergences as \( \bar{l} \to 0 \), we have to introduce a length scale \( \rho \) defined by \( \bar{l} = \epsilon \rho \) and add a counter term proportional to
\( \log \epsilon \) to cancel the divergence as \( \epsilon \) tends to zero. The counter term is

\[
S_3 = -\frac{\bar{\mathcal{l}}^3}{64\pi G} \log \epsilon \int d^4x \sqrt{\gamma} \left( \gamma^{ik} \gamma^{jl} R_{ij} R_{kl} - \frac{1}{3} R^2 \right)
\]

\[
= -\frac{\bar{\mathcal{l}}^3}{64\pi G} \log \epsilon \int d^4x \sqrt{\gamma} (-12 \\
+ \frac{1}{R^4} [2\delta^{ij} h_{ij} - \frac{3}{2} \nu^2 h_{ij} + \frac{1}{4} \delta^{ij} \nabla^2 h_{ij}]).
\]

This term does indeed cancel the logarithmic divergence, leaving us with

\[
W_{CFT} = \frac{3N^2 \Omega_4}{8\pi^2} \log \frac{R}{\rho} + (4.158) \]

\[
= \frac{N^2}{256\pi^2 R^4} \sum_p \left( \int d^4x' \sqrt{\gamma} h^{kl}(x') H_{kl}^{(p)}(x') \right)^2 \times \\
(2p(p+1)(p+2)(p+3) \log(\rho/R) + \Psi(p)).
\]

Note that varying \( W_{CFT} \) twice with respect to \( h_{ij} \) yields the expression for the transverse traceless part of the correlator \( \langle T_{ij}(x) T_{ij}(x') \rangle \) on a round four sphere. At large \( p \), this behaves like \( p^4 \log p \), as expected from the flat space result [Gubser et al. (1998)]. In fact this correlator can be determined in closed form solely from the trace anomaly and symmetry considerations. However, we shall be interested in calculating cosmologically observable effects, for which our mode expansion is more useful.

4.3.9.3 The Total Action

Recall that our 5D action is

\[
S = S_{EH} + S_{GH} + 2S_1 + W_{CFT}.
\]

In order to calculate correlators of the metric, we need to evaluate the path integral [Hawking et al. (2000)]

\[
Z[h] = \int_{B_1 \cup B_2} \mathcal{D}[\delta g] \exp(-S) = \\
\exp(-2S_1[h_0 + h] - W_{CFT}[h_0 + h]) \times \\
\left( \int_B \mathcal{D}[\delta g] \exp(-S_{EH}[g_0 + \delta g] - S_{GH}[g_0 + \delta g]) \right)^2.
\]
Here $g_0$ and $h_0$ refer to the unperturbed background metrics in the bulk and on the wall respectively and $h$ denotes the metric perturbation on the wall. Many of the terms required here can be obtained from above results by simply replacing $\bar{l}$ and $\bar{G}$ with $l$ and $G$. For example, from equation (4.157) we get

$$S_1[h_0 + h] = \frac{3l^3}{8\pi G} \int d^4x \sqrt{\hat{g}} \left( \sinh^4 y_0 - \frac{1}{4l^4} \right),$$

where $y_0$ is defined by $R = l \sinh y_0$. The path integral over $\delta g$ is performed by splitting it into a classical and quantum part:

$$\delta g = h + h',$$

where the boundary perturbation $h$ is extended into the bulk using the linearized Einstein equations and the requirement of finite Euclidean action, i.e., $h$ is given in the bulk by equation (4.153). $h'$ denotes a quantum fluctuation that vanishes at the domain wall. The gravitational action splits into separate contributions from the classical and quantum parts:

$$S_{EH} + S_{GH} = S_0[h] + S'[h'],$$

where $S_0$ can be read off from equations (4.155) and (4.156) as

$$S_0 = -\frac{3l^3\Omega_4}{2\pi G} \int_0^{y_0} dy \sinh^2 y_0 \cosh^2 y_0$$

$$+ \frac{l^3}{16\pi G} \int d^4x \sqrt{\hat{g}} \left( \frac{1}{4l^4} h^{ij} \partial_i h_{ij} + \frac{\coth y_0}{l^4} h^{ij} h_{ij} \right).$$

Note that $S'$ cannot be converted to a surface term since $h'$ does not satisfy the Einstein equations. We shall not need the explicit form for $S'$ since the path integral over $h'$ just contributes a factor of some determinant $Z_0$ to $Z[h]$. We get

$$Z[h] = Z_0 \exp(-2S_0[h_0 + h] - 2S_1[h_0 + h] - W_{CFT}[h_0 + h]).$$
The exponent is given by

\[ 2S_0 + 2S_1 + W_{\text{CFT}} = \frac{3l^3 \Omega_4}{\pi G} \int_0^{y_0} dy \sinh^2 y \cosh^2 y + \frac{3\Omega_4 R^4}{4\pi G l} \]

\[ + \frac{3N^2 \Omega_4}{8\pi^2} \log \frac{R}{\rho} + \frac{1}{l^2} \sum_p \left( \int d^4 x' \sqrt{-g} h^{ij}(x') H^{ij}_{kl}(x') \right)^2 \times \]

\[ \left[ \frac{l^3}{32\pi G} \left( \frac{f_p(y_0)}{f_p(y_0)} + 4 \coth y_0 - 6 \right) \right. \]

\[ \left. + \frac{N^2}{256\pi^2 \sinh^3 y_0} (2p(p + 1)(p + 2)(p + 3) \log(\rho/R) + \Psi(p)) \right]. \]

We have kept the unperturbed action in order to demonstrate how the conformal anomaly arises: it is simply the coefficient of the \( \log(R/\rho) \) term divided by the area \( \Omega_4 R^4 \) of the sphere. If we set the metric perturbation to zero and vary \( R \) (using \( R = l \sinh y_0 \)) then we reproduce equation (4.148).

Having calculated \( R \), we can now choose a convenient value for the renormalization scale \( \rho \). If we were dealing purely with the CFT then we could keep \( \rho \) arbitrary. However, since the third counter term involves the square of the Weyl tensor (the integrand is proportional to the difference of the Euler density and the square of the Weyl tensor), we can expect pathologies to arise if this term is present when we couple the CFT to gravity. In other words, when coupled to gravity, different choices of \( \rho \) lead to different theories. We shall choose the value \( \rho = R \) so that the third counter term exactly cancels the divergence in the CFT, with no finite remainder and hence no residual curvature squared terms in the action.

The (Euclidean) graviton correlator can be read off from the action as [Hawking et al. (2000)]

\[ \langle h_{ij}(x) h_{ij'}(x') \rangle = \frac{128\pi^2 R^4}{N^2} \sum_{p=2}^{\infty} W_{ij'j'}^{(p)}(x, x') F(p, y_0)^{-1} \] (4.159)

where we have eliminated \( l^3/G \) using equation (4.148). The function \( F(p, y_0) \) is given by

\[ F(p, y_0) = e^{y_0} \sinh y_0 \left( \frac{f_p(y_0)}{f_p(y_0)} + 4 \coth y_0 - 6 \right) + \Psi(p), \]
and the bitensor $W^{(p)}_{i j i' j'}(x, x')$ is defined as

$$W^{(p)}_{i j i' j'}(x, x') = \sum_{k, l, m, \ldots} H^{(p)}_{i} H^{(p)}_{i'}(x),$$

with the sum running over all the suppressed labels $k, l, m, \ldots$ of the tensor harmonics.

The appearance of $N^2$ in the denominator in equation (4.159) suggests that the CFT suppresses metric perturbations on all scales. This is misleading because $R$ also depends on $N$. The function $F(p, y_0)$ has the following limiting forms for large and small radius:

$$\lim_{y_0 \to \infty} F(p, y_0) = \Psi(p) + p^2 + 3p + 6, \quad (4.160)$$

$$\lim_{y_0 \to 0} F(p, y_0) = \Psi(p) + p + 6. \quad (4.161)$$

$F(p, y_0)$ has poles at $p = -4, -5, -6, \ldots$ with zeros between each pair of negative integers starting at $-3, -4$. When we analytically continue to Lorentzian signature, we shall be particularly interested in zeros lying in the range $p \geq -3/2$. There is one such zero exactly at $p = 0$, another near $p = 0$ and a third near $p = -3/2$. For large radius, these extra zeros are at $p \approx -0.054$ and $p \approx -1.48$ while for small radius they are at $p \approx 0.094$ and $p \approx -1.60$. For intermediate radius they lie between these values, with the zeros crossing through $-3/2$ and 0 at $y_0 \approx 0.632$ and $y_0 \approx 1.32$ respectively.

4.3.9.4 Comparison with 4D Gravity

Above we have discussed how the RS scenario reproduces the predictions of 4D gravity when the effects of matter on the domain wall dominates the effects of the RS-CFT. In our case we have a CFT on the domain wall. This has action proportional to $N^2$. The RS-CFT is a similar CFT (but with a cut-off) and therefore has action proportional to $N_{RS}^2$. Hence we can neglect it when $N \gg N_{RS}$. The logarithmic counterterm is also proportional to $N_{RS}^2$ and therefore also negligible. We therefore expect the predictions of 4D gravity to be recovered when $N \gg N_{RS}$. We shall now demonstrate this explicitly.

First consider the radius $R$ of the domain wall given by equation (4.148). It is convenient to write this in terms of the rank $N_{RS}$ of the RS-CFT (given
Quantum Leap

by $l^3/G = 2N_{RS}^2/\pi$

$$R^3 \sqrt{\frac{R^2}{l^2}} + 1 = \frac{N^2}{16N_{RS}} + \frac{R^4}{l^4}. $$

If we assume $N \gg N_{RS} \gg 1$ then the solution is

$$\frac{R}{l} = \frac{N}{2\sqrt{2}N_{RS}} \left[ 1 + \frac{N_{RS}^2}{N^2} + O\left(\frac{N_{RS}^4}{N^4}\right) \right].$$

Note that this implies $R \gg l$, i.e., the domain wall is large compared with the anti–de Sitter length scale.

Now let’s turn to a 4D description in which we are considering a four sphere with no interior. The only matter present is the CFT. The metric is simply

$$ds^2 = R_4^2\tilde{\gamma}_{ij}dx^i dx^j,$$

where $R_4$ remains to be determined. The action is the 4D Einstein–Hilbert action (without cosmological constant) together with $W_{CFT}$. There is no Gibbons–Hawking term because there is no boundary. Without a metric perturbation, the action is simply [Hawking et al. (2000)]

$$S = -\frac{1}{16\pi G_4} \int d^4x \sqrt{-\gamma} R + W_{CFT} = -\frac{3\Omega_4 R_4^2}{4\pi G_4} + \frac{3N^2\Omega_4}{8\pi^2} \log \frac{R_4}{\rho}$$

where $G_4$ is the 4D Newton constant. We want to calculate the value of $R_4$ so we can’t choose $\rho = R_4$ yet. Varying $R_4$ gives

$$R_4^2 = \frac{N^2 G_4}{4\pi},$$

and $N$ is large hence $R_4$ is much greater than the 4D Planck length. Substituting $G_4 = G_5/l$, this reproduces the leading order value for $R$ found above from the 5D calculation.

We can now go further and include the metric perturbation. The perturbed 4D Einstein–Hilbert action is

$$S_{EH}^{(4)} = -\frac{1}{16\pi G_4} \int d^4x \sqrt{-\gamma} \left( 12R_4^2 - \frac{2}{R_4^2} h^{ij} h_{ij} + \frac{1}{4R_4^2} \nabla^2 h_{ij} \right). \quad (4.162)$$
Adding the perturbed CFT gives

\[ S = -\frac{3N^2\Omega_4}{16\pi^2} + \frac{3N^2\Omega_4}{8\pi^2} \log \frac{R_4}{\rho} + \sum_p \left( \int d^4x' \sqrt{\gamma} h^{kl}(x') H_{kl}^{(p)}(x') \right)^2 \]

\[ \left[ \frac{1}{64\pi G_4 R_4^2} (p^2 + 3p + 6) + \frac{N^2}{256\pi^2 R_4^2} (2p(p + 1)(p + 2)(p + 3) \log(p/R_4) + \Psi(p)) \right]. \]

Setting \( \rho = R_4 \), we find that the graviton correlator for a 4D universe containing the CFT is

\[ \langle h_{ij}(x)h_{i'j'}(x') \rangle = 8N^2G_4^2 \sum_{p=2}^{\infty} W_{ij'i'j'}^{(p)}(x,x') \left[ p^2 + 3p + 6 + \Psi(p) \right]^{-1}. \]

This can be compared with the expression obtained from the 5D calculation, which can be written

\[ \langle h_{ij}(x)h_{i'j'}(x') \rangle = \frac{8N^2G^2}{l^2} \left[ 1 + \mathcal{O}(N_{RS}^2/N^2) \right] \times \sum_{p=2}^{\infty} W_{ij'i'j'}^{(p)}(x,x') \left[ p^2 + 3p + 6 + \Psi(p) \right]^{-1} + 4p(p + 1)(p + 2)(p + 3)(N_{RS}^2/N^2) \log(N_{RS}/N) + \mathcal{O}(N_{RS}^2/N^2) \left[ \right. \left. \right]^{-1}. \]

We have expanded in terms of

\[ \frac{N_{RS}^2}{N^2} = \frac{\pi l^3}{2N^2G}. \]

The four and 5D expressions clearly agree (for \( G_4 = G/l \)) when \( N \gg N_{RS} \), i.e., \( R \gg l \). There are corrections of order \( (N_{RS}^2/N^2) \log(N_{RS}/N) \) coming from the RS–CFT and the logarithmic counter term. In fact, these corrections can be absorbed into the renormalization of the CFT on the domain wall if, instead of choosing \( \rho = R \), we choose

\[ \rho = R \left( 1 - \frac{2N_{RS}^2}{N^2} \log(N_{RS}/N) \right). \]

The corrections to the 4D expression are then of order \( N_{RS}^2/N^2 \). We shall not give these correction terms explicitly although they are easily obtained from the exact result \( (4.159) \).
Lorentzian Correlator

Here we will show how the Euclidean correlator calculated above is analytically continued to give a correlator for Lorentzian signature [Hawking et al. (2000)]. Let us first introduce a new label \( p' = i(p + 3/2) \), so that on the four sphere [Hawking et al. (2000)]

\[
\hat{\nabla}^2 H^{(p')}_{ij} = \lambda_{p'} H^{(p')}_{ij}, \quad \text{where} \quad p' = 7i/2, 9i/2, \ldots \quad \text{and} \quad \lambda_{p'} = (p'^2 + 17/4).
\]

Recall that there are extra labels on the tensor harmonics that we have suppressed. The set of rank–two tensor eigenmodes on \( S^4 \) forms a representation of the symmetry group of the manifold. Hence the sum of the degenerate eigenfunctions with eigenvalue \( \lambda_{p'} \) defines a maximally symmetric bitensor \( W_{i'j'}(\mu(\Omega, \Omega')) \), where \( \mu(\Omega, \Omega') \) is the distance along the shortest geodesic between the points with polar angles \( \Omega \) and \( \Omega' \).

The motivation for the unusual labelling is that in terms of the label \( p' \) the bitensor on \( S^4 \) has exactly the same formal expression as the corresponding bitensor on Lorentzian de Sitter space. This property will enable us to analytically continue the Euclidean correlator into the Lorentzian region without Fourier decomposing it. In other words, instead of imposing by hand a prescription for the vacuum state of the graviton on each mode separately and propagating the individual modes into the Lorentzian region, we compute the two–point tensor correlator in real space, directly from the no boundary path integral. Since the path integral unambiguously specifies the allowed fluctuation modes as those which vanish at the origin of the instanton, this automatically gives a unique Euclidean correlator. The technical advantage of our method is that dealing directly with the real space correlator makes the derivation independent of the gauge ambiguities involved in the mode decomposition [Hawking et al. (2000)].

We begin by continuing the graviton correlator (equation (4.159)) obtained via the 5D calculation. The analytic continuation of the correlator for 4D gravity (equation (4.163)) is completely analogous. In terms of the new label \( p' \), the Euclidean correlator (4.159) between two points on the wall is given by [Hawking et al. (2000)]

\[
\langle h_{ij}(\Omega) h_{i'j'}(\Omega') \rangle = \frac{128\pi^2 R^4}{N^2} \sum_{p'=7i/2}^{\infty} W^{(p')}_{ij;i'j'}(\mu) G(p', y_0)^{-1}, \quad \text{where}
\]

\[(4.164)\]
\[ G(p', y_0) = F(-ip' - 3/2, y_0) = e^{y_0} \sinh y_0 \left( \frac{g_p'(y_0)}{g_p'(y_0)} + 4 \coth y_0 - 6 \right) + (p'^4 - 4ip'^3 + p'^2/2 - 5ip' - 63/16) + (p'^2 + 1/4)(p'^2 + 9/4)\psi(-ip'/2 + 5/4) + \psi(-ip'/2 + 7/4) - \psi(1) - \psi(2)) \]
with \[ g_p(y) = Q^2 - ip' - 1/2 (\coth y) \]
which follows from (4.152). The function \( G(p', y_0) \) is real and positive for all values of \( p' \) in the sum and for arbitrary \( y_0 \geq 0 \).

We have the Euclidean correlator defined as an infinite sum. However, the eigenspace of the Laplacian on de Sitter space suggests that the Lorentzian propagator is most naturally expressed as an integral over real \( p' \). We must therefore first analytically continue our result from imaginary to real \( p' \). The coefficient \( G(p', y_0) \) of the bitensor is analytic in the upper half complex \( p' \)-plane, apart from three simple poles on the imaginary axis. One of them is always at \( p' = 3i/2 \), regardless of the radius of the sphere. Let the position of the remaining two poles be written \( p_k' = i\Lambda_k(y_0) \). If we take the radius of the domain wall to be large compared with the AdS scale (which is necessary for corrections to 4D Einstein gravity to be small) then \( 0 < \Lambda_k \leq 3/2 \), with \( \Lambda_1 \sim 0 \) and \( \Lambda_2 \sim 3/2 \). Since \( G(p', y_0) \) is real on the imaginary \( p' \)-axis, the residues at these poles are purely imaginary. In order to extend the correlator into the complex \( p' \)-plane, we must also understand the continuation of the bitensor itself. The condition of regularity at opposite points on the four sphere imposed by the completeness relation is sufficient to uniquely specify the analytic continuation of \( W^{(p')_{ij\pi \pi'}}(\mu) \) into the complex \( p' \)-plane.

Now we are able to write the sum in equation (4.164) as an integral along a contour \( C_1 \) encircling the points \( p' = 7i/2, 9i/2, ..., ni/2 \), where \( n \) tends to infinity. This yields
\[
\langle h_{ij}(\Omega)h_{i'j'}(\Omega') \rangle = \frac{-i64\pi^2 R^4}{N^2} \int_{C_1} dp' \tanh p' \pi W^{(p')_{ij\pi \pi'}}(\mu)G(p', y_0)^{-1}.
\]
\[(4.165)\]

Since we know the analytic properties of the integrand in the upper

\[51\] If we decrease the radius of the domain wall, then the poles move away from each other. Their behavior follows from the discussion below equations (1.160) and (1.161). For \( y_0 \leq 0.632 \), \( \Lambda_1 \) becomes slightly smaller than zero while for \( y_0 \leq 1.32 \), \( \Lambda_2 \) becomes slightly greater than \( 3/2 \).
half complex $p'$--plane, we can distort the contour for the $p'$ integral to run along the real axis. At large imaginary $p'$ the integrand decays and the contribution vanishes in the large $n$ limit. However as we deform the contour towards the real axis, we encounter three extra poles in the $\cosh p'\pi$ factor, the pole at $p' = 3i/2$ becoming a double pole due to the simple zero of $G(p', y_0)$. In addition, we have to take into account the two poles of $G(p', y_0)^{-1}$ at $p' = i\Lambda_k$.

For the $p' = 5i/2$ pole, it follows from the normalization of the tensor harmonics that $W_{ij'j'}(5i/2) = 0$. Indirectly, this is a consequence of the fact that spin--2 perturbations do not have a dipole or monopole component. The meaning of the remaining two poles of the $\tanh p'\pi$ factor has been extensively discussed in [Hawking et al. (2000)], where the continuation is described of the two--point tensor fluctuation correlator from a 4D--$O(5)$ instanton into open de Sitter space. They represent non--physical contributions to the graviton propagator, arising from the different nature of tensor harmonics on $S^4$ and on Lorentzian de Sitter space. In fact, a degeneracy appears between $p'_t = 3i/2$ and $p'_v = i/2$ tensor harmonics and respectively $p'_s = 5i/2$ vector harmonics and $p'_s = 5i/2$ scalar harmonics on $S^4$. More precisely, the tensor harmonics that constitute the bitensors $W_{ij'j'}(3i/2)$ and $W_{ij'j'}(i/2)$ can be constructed from a vector (scalar) quantity. Consequently, the contribution to the correlator from the former pole is pure gauge, while the latter eigenmode should really be treated as a scalar perturbation, using the perturbed scalar action. Henceforth we shall exclude them from the tensor spectrum. This leaves us with the poles of $G(p', y_0)$ at $p' = i\Lambda_k$.

If we deform the contour towards the real axis, we must compensate for them by subtracting their residues from the integral. We will see that these residues correspond to discrete ‘super–curvature’ modes in the Lorentzian tensor correlator.

The contribution from the closing of the contour in the upper half $p'$--plane vanishes. Hence our final result for the Euclidean correlator reads [Hawking et al. (2000)]

$$
\langle h_{ij}(\Omega)h_{ij'}(\Omega') \rangle = \frac{-i64\pi^2 R^4}{N^2} \left[ \int_{-\infty}^{+\infty} dp' \tanh p'\pi W_{ij'j'}(p') G(p', y_0)^{-1}
+ 2\pi \sum_{k=1}^{2} \tan \Lambda_k \pi W_{ij'j'}^{(i\Lambda_k)}(\mu) \text{Res}(G(p', y_0)^{-1}; i\Lambda_k) \right].
$$

The analytic continuation from a four sphere into Lorentzian closed de Sitter space is given by setting the polar angle $\Omega = \pi/2 - it$. Without loss
of generality we may take $\mu = \Omega$, and $\mu$ then continues to $\pi/2 - it$. We then get the correlator in de Sitter space where one point has been chosen as the origin of the time coordinate.

The extra factor $ie^{ip\pi}/\sinh p'\pi$ combines with the factor $-it\tanh p'\pi$ in the integrand to $\cosh p'\pi$. Furthermore, since $G(-p', y_0) = G(p', y_0)$, we can rewrite the correlator as an integral from 0 to $\infty$. We finally get the Lorentzian tensor Feynman (time–ordered) correlator,

$$\langle h_{ij}(x)h_{i'j'}(x') \rangle = \frac{128\pi^2 R^4}{N^2} \left[ \int_0^{+\infty} dp' \tanh p'\pi W^{L(i\Lambda_k)}_{ij'j'}(\mu) \Re(G(p', y_0)^{-1}) + \pi \sum_{k=1}^2 \tan \Lambda_k \pi W^{L(i\Lambda_k)}_{ij'j'}(\mu) \Re(G(p', y_0)^{-1}; i\Lambda_k) \right] + \frac{i28\pi^2 R^4}{N^2} \left[ \int_0^{+\infty} dp' W^{L(p')}_{ij'j'}(\mu) \Re(G(p', y_0)^{-1}) + \pi \sum_{k=1}^2 W^{L(i\Lambda_k)}_{ij'j'}(\mu) \Re(G(p', y_0)^{-1}; i\Lambda_k) \right].$$

In this integral the bitensor $W^{L(p')}_{ij'j'}(\mu(x, x'))$ may be written as the sum of the degenerate rank–two tensor harmonics on closed de Sitter space with eigenvalue $\lambda_{p'} = (p'^2 + 17/4)$ of the Laplacian. Note that the normalization factor $\tilde{Q}_{p'} = p'(4p'^2 + 25)/48\pi^2$ of the bitensor is imaginary at $p' = i\Lambda_k$ and the residues of $G^{-1}$ are also imaginary, so the quantities in square brackets are all real. Both integrands in this equation vanish as $p' \to 0$, so the correlator is well-behaved in the infrared.

For cosmological applications, one is usually interested in the expectation of some quantity squared, like the microwave background multipole moments. For this purpose, all that matters is the symmetrized correlator, which is just the real part of the Feynman correlator.

Gravitational waves provide an extra source of time–dependence in the background in which the cosmic microwave background photons propagate. In particular, the contribution of gravitational waves to the CMB anisotropy is given by the integral in the Sachs–Wolfe formula, which is basically the integral along the photon trajectory of the time derivative of the tensor perturbation. Hence the resulting microwave multipole moments $C_l$ can be directly determined from the graviton correlator.

We can therefore understand the effect of the strongly coupled CFT on the microwave fluctuation spectrum. On the 4–sphere, this is easily obtained by varying the Einstein–Hilbert action with a cosmological constant.
Quantum Leap

In terms of the bitensor, this yields [Hawking et al. (2000)]

\[
\langle h_{ij}(\Omega)h_{i'j'}(\Omega') \rangle = 32\pi G_4 R^2 \sum_{p' = 7i/2}^{\infty} W_{ijij'}(\mu(\Omega, \Omega')) \frac{\lambda_{p'} - 2}{\lambda_{p'}},
\]

which continues to

\[
\langle h_{ij}(x)h_{i'j'}(x') \rangle = 32\pi G_4 R^2 \int_{0}^{+\infty} dp' \frac{d^4 p'}{\lambda_{p'} - 2} W_{ijij'}(\mu(x, x')).
\]

Note that (apart from the pole at $p' = 3i/2$ corresponding to the gauge mode mentioned before) there are no supercurvature modes.

4.3.9.5 Summary on Hawking’s Brane New World

We have studied a Randall–Sundrum cosmological scenario consisting of a domain wall in anti–de Sitter space with a large $N$ conformal field theory living on the wall. The conformal anomaly of the CFT provides an effective tension which leads to a de Sitter geometry for the domain wall.

We have computed the spectrum of quantum mechanical vacuum fluctuations of the graviton field on the domain wall, according to Euclidean no boundary initial conditions. The Euclidean path integral unambiguously specifies the tensor correlator with no additional assumptions. This is the first calculation of quantum fluctuations for RS cosmology.

In the usual inflationary models, one considers the classical action for a single scalar field. In that context, it is consistent to neglect quantum matter loops, on the grounds that they are small. On the other hand, in this section we have studied a strongly coupled large $N$ CFT living on the domain wall, for which quantum loops of matter are important. By using the AdS/CFT correspondence, we have performed a fully quantum mechanical treatment of this CFT. The most notable effect of the large $N$ CFT on the tensor spectrum is that it suppresses small scale fluctuations on the microwave sky. It can be seen that the CFT yields a $(p'^4 \ln p'^{-2})$ behavior for the graviton propagator at large $p'$ (in agreement with the flat space results of Tomboulis [1977]), instead of the usual $p'^{-2}$ falloff.

In other words, quantum loops of the CFT give space–time a rigidity that strongly suppresses metric fluctuations on small scales. Note that this is true independently of how the de Sitter geometry arises, i.e., it is also true for 4D Einstein gravity. In addition, the coupling of the CFT to tensor perturbations gives rise to two additional discrete modes in the tensor spectrum. Although this is a novel feature in the context of inflationary
tensor perturbations, it is not surprising. In conventional open inflationary scenarios for instance, the coupling of scalar field fluctuations with scalar metric perturbations introduces a supercurvature mode with an eigenvalue of the Laplacian close to the discrete de Sitter gauge mode \cite{Yamamoto et al. (1996)}. The former discrete mode at $p' = i\Lambda_1 \sim 3i/2$ is the analogue of this well known supercurvature mode in the scalar fluctuation spectrum. The second mode has an eigenvalue $p' = i\Lambda_2 \sim 0$. Its interpretation is less clear, but it is clearly an effect of the matter on the domain wall. However it hardly contributes to the correlator because $\tan \Lambda_2 \pi$ is very small.

The effect of the CFT on large scales is more difficult to quantify because of the complicated $p'$—dependence of the tensor correlator in the low—$p'$ regime. Generally speaking, however, long—wavelength tensor correlations in closed (or open) models for inflation are very sensitive to the details of the underlying theory, as well as to the boundary conditions at the instanton. Since tensor fluctuations do give a substantial contribution to the large scale CMB anisotropies, this may provide an additional way to observationally distinguish different inflationary scenarios.

Most matter fields can be expected to behave like a CFT at small scales. Furthermore, fundamental theories such as string theory predict the existence of a large number of matter fields. Therefore, our results based on a quantum treatment of a large $N$ CFT may be accurate at small scales for any matter. If this is the case then our result shows that tensor perturbations at small angular scales are much smaller than predicted by calculations that neglect quantum effects of matter fields.

### 4.3.10 Brane—World Quantum Cosmology

Recall that in the brane-world picture, branes are time-like surfaces with metric $g_{\alpha\beta}$ embedded into bulk space-time with metric $G_{AB}$. The action functional can be taken to be the sum of a 4D (brane) and 5D (bulk) contribution \cite{Barvinsky and Nesterov (2006); Esposito (2006)}.

$$S = S_4[g_{\alpha\beta}(x)] + S_5[G_{AB}(X)].$$

In general, there exist vector fields $R_B, R_\nu$ on the space of histories such that

$$R_B S_5 = 0, \quad R_\nu S_4 = 0,$$
with Lie brackets given by

\[ [R_B, R_D] = C^A_{BD} R_A, \quad [R_\mu, R_\nu] = C^\lambda_{\mu\nu} R_\lambda. \]

The components of the vector fields \( R_B \) and \( R_\nu \) generate 5D and 4D diffeomorphisms, respectively, while the bulk and brane ghost operators read (with \( F^A \) and \( \chi^\mu \) the bulk and brane gauge–fixing functionals, respectively)

\[ Q^A_B \equiv R_B F^A = F^A_{\alpha} R^\alpha_B, \]
\[ J^\nu_\nu \equiv R_\nu \chi^\mu = \chi^\mu_i R^i_\nu. \]

On denoting by \( S_A \) and \( T^B \) the bulk ghost fields, the cosmological wave function of the bulk space-time can be written as [Barvinsky and Nesterov (2006)]

\[ \psi_{\text{Bulk}} = \int_{G_{AB}[\partial M] = g_{\alpha\beta}} \mu(G_{AB}, S, T) e^{i\bar{S}_5}, \]

with \( \mu(G_{AB}, S, T) \) a suitable measure functional, while

\[ \bar{S}_5 = S_5[G_{AB}] + \frac{1}{2} F^A \omega_{AB} F^B + S_A Q^A_B T^B. \]

The brane–world effective action \( \Gamma \) can (in principle) be obtained from the formula [Barvinsky and Nesterov (2006); Esposito (2006)]

\[ e^{i\Gamma} = \int \nu(g_{\alpha\beta}, \rho, \sigma) e^{i\bar{S}_4} \psi_{\text{Bulk}}, \]

where \( \nu(g_{\alpha\beta}, \rho, \sigma) \) is a suitable measure functional over brane metrics and brane ghost fields, while

\[ \bar{S}_4 = S_4 + \frac{1}{2} \chi^\mu C_{\mu\nu} \chi^\nu + \rho_\mu J^\mu_\nu \sigma^\nu. \]

Recent developments in this respect can be found in [Barvinsky and Nesterov (2006)], where the authors lay the foundations for a systematic application of the background–field method to the brane–world picture.
Cosmic Inner Flight

We will finish this chapter on cosmic dynamics by quoting the words of Paramhansa Yogananda, depicting the echo of the Big–Bang occurring 13.7 billion years ago:

“My body became immovably rooted; breath was drawn out of my lungs as if by some huge magnet. Soul and mind instantly lost their physical bondage, and streamed out like a fluid piercing light from my every pore. The flesh was as though dead, yet in my intense awareness I knew that never before had I been fully alive. My sense of identity was no longer narrowly confined to a body, but embraced the circumambient atoms. People on distant streets seemed to be moving gently over my own remote periphery... The whole vicinity lay bare before me. My ordinary frontal vision was now changed to a vast spherical sight, simultaneously all-perceptive... All objects within my panoramic gaze trembled and vibrated like quick motion pictures... The unifying light alternated with materializations of form, the metamorphoses revealing the law of cause and effect in creation... An oceanic joy broke upon calm endless shores of my soul. The Spirit of God, I realized, is exhaustless Bliss; His body is countless tissues of light. A swelling glory within me began to envelop towns, continents, the earth, solar and stellar systems, tenuous nebulae, and floating universes. The entire cosmos, gently luminous, like a city seen afar at night, glimmered within the infinitude of my being. The sharply etched global outlines faded somewhat at the farthest edges; there I could see a mellow radiance, ever–undiminished. It was indescribably subtle; the planetary pictures were formed of a grosser light... The divine dispersion of rays poured from an Eternal Source, blazing into galaxies, transfigured with ineffable auras. Again and again I saw the creative beams condense into constellations, then resolve into sheets of transparent flame. By rhythmic reversion, sextillion worlds passed into diaphanous luster; fire became firmament. I cognized the center of the empyrean as a point of intuitive perception in my heart. Irradiating splendor issued from my nucleus to every part of the universal structure. Blissful amrita, the nectar of immortality, pulsed through me with a quick silver–like fluidity. The creative voice of God I heard resounding as AUM, the vibration of the Cosmic Motor.”
This page intentionally left blank
Chapter 5

Quantum Biophysics in Human Body: Electro–Muscular Stimulation

In this chapter we develop covariant biophysics of electro–muscular stimulation, as an externally induced generator of our covariant muscular forces, $F_i = mg_{ij}a^j$ (see [Ivancevic and Ivancevic (2007b)]). The so–called functional electrical stimulation (FES) of human skeletal muscles is used in rehabilitation and in medical orthotics to externally stimulate the muscles with damaged neural control (see [Ivancevic and Ivancevic (2005)]). However, the repetitive use of electro–muscular stimulation, besides functional, causes also structural changes in the stimulated muscles, giving the physiological effect of muscular training.

5.1 Basics of Electrical Muscular Stimulation

The use of low and very low frequency impulses in the body, delivered through electrodes, is known as transcutaneous stimulation of the nerves, electro–acupuncture and electro–stimulation. Here, an electromagnetic field accompanies the passage of the electric current through the conductive wire. This is generally known as the term ‘electromagnetic therapy’.

In the original sense acupuncture meant the inserting of needles in specific regions of the body. Electro–acupuncture supplies the body with low–volt impulses through the medium of surface electrodes to specific body regions or by non specific electrodes. Transcutaneous electric stimulation of the nerves (TENS) has for years been a well known procedure in conventional medicine. The impulses that are produced with this type of stimulation, are almost identical with those of electro–stimulation, yet many doctors still assume, that they are two different therapies. This has resulted in TENS being considered as a daily therapy, while electro–acupuncture or electro–stimulation were treated as ‘alternative therapy’. Apart from
the fact, that electro–acupuncture electro–impulses are delivered through needles, both therapies should be considered identical. Patients, who have reservations about the use of needles, can by the use of electric impulses over surface electrodes on the skin, have a satisfactory alternative (see Figure 5.1). We choose the term electro–stimulation, because the acupuncture system is not included in all therapies. Clinical tests have showed that there are two specific types of reactions:

- The first reaction is spontaneous and dependent on the choice of body region. The stimulation of this part of the body results in an unloading, that can be compared with that of a battery. Normally this goes hand in hand with an immediate improvement in the patient. This effect of unloading may also be reached by non–specific electric stimulation.
- The second normal reaction is of a delayed nature, that results in relaxation and control of pain. Moreover two other important effects follow, that begin between 10 and 20 minutes after the start of the treatment. This reaction is associated (combined) with different chemicals, such as beta–endorphins and 5–hydroxytryptamins. When using low and very low frequency stimulations, the second effect is obtained by the utilization of specific frequencies on the body. This is independent of

---

Fig. 5.1 Schematic of electrical muscular stimulation \( EMS \).
the choice of a specific part of the body, because the connected electromagnet makes the induction of secondary electric current in the whole body possible.

Now, when cosmetic surgeons perform electrical muscular stimulation (EMS, for short) on the human face or body (as schematically depicted on Figure 5.1), they usually take for granted half-a-dozen biophysical processes that are actually involved in this apparently simple stimulus–response–type action.

When the surface electrical muscular stimulation EMS pads are applied to the certain place of the human face or body, the first considerable tissue reaction is depolarization of sarcolemma, close to the electrodes. Muscular ability (similar to the neural one, but about ten times slower) to produce an action potential as a response to the stimulation, is termed ‘excitability’. By means of the EMS, current is passed across a membrane to produce a transient depolarization of the resting potential of sarcolemma, which, if it is of sufficient duration and magnitude, can initiate the train of events
that produces muscular action potential (see Figure [5.1]). The minimum necessary intensity of stimulus is called the threshold stimulus. The term ‘threshold’ is commonly used to refer either to the absolute magnitude of the muscle–cell membrane potential at which an action potential is initiated or to the magnitude of depolarization from resting potential (in which the membrane naturally polarizes sodium and potassium ions) required to initiate an action potential. A stimulus of less than threshold intensity is referred to as subthreshold, one of greater than threshold intensity as super–threshold. The threshold potential for excitation is not a fixed parameter. The thresholds of different muscle–membranes may vary considerably. Furthermore, the threshold of a single cell can change, either rapidly, as after a train of impulses produced by the $EMS_{total}$, or more slowly, in response to metabolic or hormonal influences.

On the other hand, in view of modern biophysics, there are six distinctive phases of electrical muscular stimulation, as usually performed using the surface pads (like those on Figures [5.3] and [5.2]):

1. **Electrodynamic stimulation fields**, consisting of:
   - (a) External Maxwell electrodynamics (smooth, causal, unique and fully predictive); and
   - (b) Internal cellular bio–quantum electrodynamics (rapidly fluctuating, uncertain and stochastic, allowing only probabilistic approach).

2. **Muscular contraction paths**, consisting of:
   - (a) Anatomical external muscular mechanics; and
   - (b) Myofibrillar internal cellular bio–quantum mechanics.

3. **Geometric face & body shapes and curvatures**, consisting of:
   - (a) Smooth 2D external skin geometry; and
   - (b) Coarse–grained and fractal, internal nD cellular muscle–fat geometry.

Combined together, these six electro–mechano–geometric faces of electro–muscular stimulation generate the three–link $EMS$–transition functor:

\[
\text{ELECTRICAL STIMUL} \Rightarrow \text{MUSCULAR CONTRACT} \Rightarrow \text{FACE or BODY SHAPE}
\]

The $EMS$ transition functor is based on Feynman–like experimental approach to electrical muscular stimulation (see Figure [5.4]): the flow of electric current from the positive surface pad $A^\oplus$ to the negative pad $B^\ominus$.
can be approximated by the vector sum of complex vectors \( \sum k \rho_k e^{i\theta_k} = \rho_k (\cos \theta_k + i \sin \theta_k) \), (in the complex plane), where \( \theta_k \) are proportional to the time taken by each vector \( \rho_k \). This vector sum will be developed into the proper Feynman path integral (see [Feynman (1998)]).

The purpose of this chapter is a modern and rigorous description of the above transition map, by elaboration of the six electro–mechano–geometric facets of the surface electro–muscular stimulation.

### 5.2 EMS Functor

Biophysically, electrical muscular stimulation represents a union of external electrical stimulation fields, internal myofibrillar excitation–contraction paths, and dissipative skin & fat geometries, formally written as

\[
EMS_{total} = EMS_{fields} \cup EMS_{paths} \cup EMS_{geom}. \tag{5.1}
\]
Following the current trends of the XXI century biophysics, corresponding to each of the three EMS–phases in 5.1 we formulate:

1. The least action principle, to model a unique, external–anatomical, predictive and smooth, macroscopic EMS field–path–geometry; and
2. Associated Feynman path integral, to model an ensemble of rapidly and stochastically fluctuating, internal, microscopic, fields–paths–geometries of the cellular EMS, to which the external–anatomical macro–level represents both time and ensemble average.\[1\]

In the proposed formalism, muscular excitation–contraction paths \(x^i(t)\) are caused by electrodynamic stimulation fields \(F^k(t)\), while they are both affected by dissipative and noisy skin & fat shapes and curvatures, defined by the local Riemannian metric tensor \(g_{ij}\).

In the following text, we first formulate the global model for the \(EMS_{total}\), to set up the general formalism to be specialized subsequently for each of the three EMS–phases.

\[\text{Fig. 5.4 Simplified Feynman–like experimental approach to electrical muscular stimulation: the flow of electric current from the positive surface pad } A^\oplus \text{ to the negative pad } B^\ominus \text{ can be approximated by the vector sum of complex vectors } \rho_k e^{i\theta_k}, \text{ where } \theta_k \text{ are proportional to the time taken by each vector } \rho_k; \text{ this vector sum will be further developed into Feynman integral.}\]

### 5.2.1 Global Macro–Level of \(EMS_{total}\)

In general, at the macroscopic EMS–level we first formulate the total action \(S[\Phi]\), our central quantity, which can be described through physical dimensions of Energy \(\times\) Time = Effort (which is also the dimension of the

\[1\]Recall that ergodic hypothesis equates time average with ensemble average.
Planck constant $\hbar$ (= 1 in normal units) (see, e.g., Deligne et al. (1999)).

This total action quantity has immediate biophysical ramifications: the greater the action – the higher the stimulation effect on the new shape. The action $S[\Phi]$ depends on macroscopic fields, paths and geometries, commonly denoted by an abstract field symbol $\Phi^i$. The action $S[\Phi]$ is formally defined as a temporal integral from the initial time instant $t_{ini}$ to the final time instant $t_{fin}$,

$$S[\Phi] = \int_{t_{ini}}^{t_{fin}} L[\Phi] dt,$$

with Lagrangian density, given by

$$L[\Phi] = \int d^n x L(\Phi^i, \partial_x \Phi^i),$$

where the integral is taken over all $n$ coordinates $x^j = x^j(t)$ of the EMS, and $\partial_x \Phi^i$ are time and space partial derivatives of the $\Phi^i$–variables over coordinates.

Second, we formulate the least action principle as a minimal variation $\delta$ of the action $S[\Phi]$

$$\delta S[\Phi] = 0,$$

which, using variational Euler–Lagrangian equations (see section 3.3 above), derives field–motion–geometry of the unique and smooth EMS–transition functor

$$T : STIMUL_{t_{ini}} \Rightarrow CONTRACT_{t_{mid}} \Rightarrow SHAPE_{t_{fin}},$$

acting at a macro–level from some initial time $t_{ini}$ to the final time $t_{fin}$ (via the intermediate time $t_{mid}$).

Here, we have in place $n$–categorical Lagrangian–field structure on the muscular Riemannian configuration manifold $M$,

$$\Phi^i : [0, 1] \to M, \Phi^i : \Phi^i_0 \mapsto \Phi^i_1,$$

using

$$\frac{d}{dt} f_{x^i} = f_{x^i} \Rightarrow \partial_\mu \left( \frac{\partial L}{\partial \dot{\Phi}^i} \right) = \frac{\partial L}{\partial \Phi^i},$$

with

$$[x_0, x_1] \mapsto [\Phi^i_0, \Phi^i_1].$$
In this way, we get macro–objects in the global $EMS$: a single electrodynamical stimulation field described by Maxwell field equations, a single muscular excitation–contraction path described by Lagrangian equation of motion, and a single Riemannian skin & fat geometry.

5.2.2 Local Micro–Level of $EMS_{total}$

After having properly defined macro–level $EMS_{total}$, with a unique and globally–smooth $EMS$–transition functor $T$, we move down to the microscopic cellular $EMS$–level of rapidly fluctuating electrodynamical fields, sarcomere–contraction paths and coarse–grained, fractal muscle–fat geometry, where we cannot define a unique and smooth field–path–geometry. The most we can do at this level of fluctuating noisy uncertainty, is to formulate an adaptive path integral and calculate overall probability amplitudes for ensembles of local transitions from negative $EMS$–pad $A^\ominus$ to the positive pad $B^\oplus$ (see Figure 5.4). This probabilistic transition micro–dynamics is given by a multi field–path–geometry, defining the microscopic transition amplitude corresponding to the macroscopic $EMS$–transition functor $T$. So, what is externally the transition functor, internally is the transition amplitude. The absolute square of the transition amplitude is the transition probability.

Now, the total $EMS$–transition amplitude, from the initial state $STIMUL$, to the final state $SHAPE$, is defined on $EMS_{total}$

$$\langle SHAPE|STIMUL\rangle_{total} = \int_{\Sigma} D[w\Phi] e^{iS[\Phi]},$$

given by modern adaptive generalization of the classical Feynman path integral, see [Feynman and Hibbs (1965)] [Feynman (1972)] [Feynman (1998)] [Deligne et al. (1999)]. The transition map (7.40) calculates overall probability amplitude along a multitude of wildly fluctuating fields, paths and geometries, performing the microscopic transition from the micro–state $STIMUL_{t_0}$ occurring at initial micro–time instant $t_0$ to the micro–state $SHAPE_{t_1}$ at some later micro–time instant $t_1$, such that all micro–time instants fit inside the global transition interval $t_0,t_1,...,t_s \in [t_{ini},t_{fin}]$. It is symbolically written as

$$\langle SHAPE|STIMUL\rangle_{total} = \int_{\Sigma} D[w\Phi] e^{iS[\Phi]},$$

where the Lebesgue integration is performed over all continuous $\Phi_{con} = fields + paths + geometries$, while summation is performed over all discrete processes and regional topologies $\Phi_{dis}$. The symbolic differential $D[w\Phi]$
in the general path integral (7.70), represents an adaptive path measure, defined as a weighted product

\[ D[w\Phi] = \lim_{N \to \infty} \prod_{s=1}^{N} w_s d\Phi_s^i, \quad (i = 1, ..., n = con + dis), \quad (5.6) \]

which is in practice satisfied with a large \( N \).

In the exponent of the path integral (7.70) we have the action \( S[\Phi] \) and the imaginary unit \( i = \sqrt{-1} \) (i can be converted into the real number \( -1 \) using the so–called Wick rotation). Feynman path integrals are usually computed by the use of perturbative expansion methods; for other non–standard applications of Feynman path integrals see [Ingber (1997); Ingber (1998)].

In this way, we get a range of micro–objects in the local \( EMS_{\text{total}} \) at the short time–level: ensembles of rapidly fluctuating, noisy and crossing electrical stimulation fields, myofibrillar contraction paths and local skin & fat shape–geometries. However, by averaging process, both in time and along ensembles of fields, paths and geometries, we can recover the corresponding global, smooth and fully predictive, external \( EMS_{\text{total}} \) transition–dynamics \( T \).

### 5.2.3 Micro–Level Adaptation and Muscular Training

The adaptive path integral (7.70) incorporates the local muscular training process according to the basic learning formula (see e.g., [Grossberg (1982), Ivancevic et al. (1999a)])

\[ \text{NEW VALUE} = \text{OLD VALUE} + \text{INNOVATION}, \]

where the term \( \text{VALUE} \) represents respectively biological images of the \( \text{STIMUL}, \text{CONTRACT} \) and \( \text{SHAPE} \).

The general synaptic weights \( w_s = w_s(t) \) in (7.71) are updated by the homeostatic neuro–muscular feedbacks during the transition process \( T \), according to one of the two standard neural training schemes, in which the micro–time level is traversed in discrete steps, i.e., if \( t = t_0, t_1, ..., t_s \) then \( t + 1 = t_1, t_2, ..., t_{s+1} \):

1. **A self–organized, unsupervised**, e.g., Hebbian–like training rule [Hebb (1949)];

   \[ w_s(t+1) = w_s(t) + \frac{\sigma}{\eta}(w_s^d(t) - w_s(t)), \quad (5.7) \]
Quantum Leap

where $\sigma = \sigma(t)$, $\eta = \eta(t)$ denote signal and noise, respectively, while superscripts $d$ and $a$ denote desired and achieved muscular micro–states, respectively; or

(2) A certain form of a supervised gradient descent training:

$$w_s(t + 1) = w_s(t) - \eta \nabla J(t), \quad (5.8)$$

where $\eta$ is a small constant, called the step size, or the training rate, and $\nabla J(n)$ denotes the gradient of the ‘performance hyper–surface’ at the $t$–th iteration.

5.3 Electrical Stimulation Fields: $EMS_{fields}$

5.3.1 External Smooth Maxwell Electrodynamics

On the macro–level in the phase $EMS_{fields}$ we formulate the electrodynamic field action principle (see, e.g. [Deligne et al. (1999)])

$$\delta S[F] = 0, \quad (5.9)$$

with the action $S[F]$ dependent on $N$ electrodynamic stimulation fields $F^i = F^i(x)$, defined as a temporal integral

$$S[F] = \int_{t_{ini}}^{t_{fin}} \mathcal{L}[F] dt, \quad (5.10)$$

with Lagrangian density given by

$$\mathcal{L}[F] = \int d^n x L(F, \partial x, F^i),$$

where the integral is taken over all $n$ coordinates $x^j = x^j(t)$ of the EMS, and $\partial x^j F^i$ are partial derivatives of the electrodynamic field variables over coordinates.

The action principle (7.53) implies the following Maxwell electrodynamics, presented here in vector, tensor and modern exterior differential form.

Given the following 3D vector–fields: the electrical field $E$, the magnetic field $B$ and the electrical current $J$, as well as the scalar electrical potential $\rho$, the Maxwell electrical vector equations read\(^2\) (see, e.g., [Misner et al. (1973)]):

\(^2\)Only electrodynamic half of the Maxwell electro–magnetic field is elaborated here, as the other, magnetodynamic part has a minor role in physiology of electro–muscular stimulation.
(1) **Electrostatics:**
\[ \nabla \cdot \vec{E} \equiv \text{div } \vec{E} = 4\pi \rho, \quad \text{and} \]

(2) **Electrodynamics:**
\[ \partial_t \vec{E} - \nabla \times \vec{B} \equiv \partial_t \vec{E} - \text{curl } \vec{B} = -4\pi \vec{J}. \]

Otherwise, given the 4D electromagnetic tensor–field **Faraday**, 
\[
F^{\alpha\beta} = \begin{pmatrix}
0 & E_x & E_y & E_z \\
-E_x & 0 & B_z & -B_y \\
-E_y & -B_z & 0 & B_x \\
-E_z & B_y & -B_x & 0
\end{pmatrix},
\]
together with the 4D **electric current** vector–field \( \vec{J}^\alpha = (\vec{J}, -\rho) \), the tensor Maxwell equation reads (with **electrostatics and electrodynamics combined**):
\[
F^{\alpha\beta}_{\alpha\beta} = 4\pi J^\alpha.
\]

Finally, given the two–form **Maxwell** \( \ast F \equiv F_{\alpha\beta} \), which is a dual of the **Faraday** tensor, (also calculated as \( F = dA \), where \( A \) is the one–form of **electrical potential**), and the three–form **charge** \( \ast J \) (which is the dual one–form), the exterior Maxwell equation reads:
\[
d\ast F = 4\pi \ast J.
\]

The two–form **Maxwell** \( \ast F \equiv F_{\alpha\beta} \), defines the **Lorentz force** one–form of the electro–muscular stimulation field,
\[
Q_\alpha \equiv \dot{p}_\alpha \propto eF_{\alpha\beta}v^\beta,
\]
where \( e \) is total electric charge and \( v^\beta \) is the velocity vector–field of the stimulation flow. This equation says that the muscular force \( Q_\alpha \) generated by the simulation is proportional to the stimulation field strength \( F_{\alpha\beta} \), velocity of the stimulation flow \( v^\beta \) through the skin–fat–muscle tissue, as well as the total stimulation charge \( e \).

Now, let \( M \) be a smooth \( n \)D closed manifold with a Riemannian metric \( g_{ij} \) and also with an exact two–form \( F = dA \). Consider the problem of existence of closed extremals of the functional
\[
S(\gamma) = \int_\gamma (\sqrt{g_{ij} \dot{x}^i \dot{x}^j} + A_i \dot{x}^i) \, dt,
\]
Quantum Leap

5.3.2 Internal Cellular Bio–Quantum Electrodynamics

At the same time, on the micro–level in the phase EMS fields we have the Feynman–type sum over fields $F_i$ ($i = 1, ..., N$) given by the adaptive path integral (see, e.g., Deligne et al. (1999))

$$ \text{STIMUL}_{\text{fields}} = \int \mathcal{D}[wF] e^{iS[F]} \text{Wick} \int \mathcal{D}[wF] e^{-S[F]}, \quad (5.11) $$

with action $S[F]$ given by the temporal integral (7.54), while Wick denotes the so–called Wick–rotation of the time variable $t$ to imaginary values $t \mapsto \tau = it$. The resulting bio–quantum field represents the bundle of cellular electrodynamic flux tubes.

Now, during the XX century, the electrodynamic flux tubes were described by the Dirac–Schwinger–Tomonaga equations of quantum electrodynamics. Today, the similar kind of flux tubes is in a more sophisticated way described by the conformal Landau–Ginzburg model (see, e.g., Deligne et al. (1999)).

Technical details of these advanced physical theories are beyond the scope of the present article. In simplified terms, we can say that they all describe field–generated solitons (see the next section for a solitary model of muscular excitation–contraction). As a result, the theory of integrable

---

3 Mathematically, quantum electrodynamics has the structure of an Abelian gauge theory with a $U(1)$ gauge group. The gauge field which mediates the interaction between the charged spin 1/2 fields is the electromagnetic field. Physically, this translates to the picture of charged particles interacting with each other by the exchange of photons.

4 In classical mathematical physics, by a soliton one usually means a “travelling wave” solution of a nonlinear partial differential equation $u_t = F(u, u_x, ...)$, i.e., a solution of the form $u(x, t) = f(x - vt)$. Solitons play a very important role in the theory of integrable mechanical systems, where any solution can be approximated by a superposition of solitons moving at different velocities, as we have in the next section where we develop a solitary model of muscular excitation–contraction. As a result, the theory of integrable
of muscular excitation–contraction). The main point of all these quantum field theories and their biophysical applications is that their macro-level averaging lift (either in time or across the ensemble of cellular tubes) produces the classical Maxwell electrodynamics \((5.3.1)\) above. On their own, they describe rapidly fluctuating, fractal and noisy, electrodynamic fields flowing from the source \((-\)) electrode to the sink \((+)\) electrode – as described in the Schwinger formalism (see, e.g., his lecture in the Nobel e–Museum).

### 5.4 Stimulated Muscular Contraction Paths: \(EMS_{paths}\)

#### 5.4.1 External Anatomical Muscular Mechanics

On the macro-level in the phase \(EMS_{paths}\) we have the muscular contraction action principle

\[
\delta S[x] = 0,
\]

with the mechanical action \(S[x]\) given by

\[
S[x] = \int_{t_{ini}}^{t_{fin}} dt \left[ \frac{m}{2} \dot{x}_i^2 + F(x) \right],
\]

where overdot denotes time derivative, so that \(\dot{x}_i\) represents the external (anatomical) muscular contraction speed, while \(m\) denotes the total estimated mass of the stimulated muscle. The corresponding Euler–Lagrangian equation, with the kinetic energy of muscular contraction

\[
E_{kin} = \frac{1}{2} g_{ij} \dot{x}_i \dot{x}_j,
\]

generated by muscular Riemannian metrics \(g_{ij}\) (see the next section on anatomical geometry), gives the Newtonian equation of motion (see e.g., [Arnold (1989)])

\[
\frac{d}{dt} \frac{\partial E_{kin}}{\partial \dot{x}_i} - \frac{\partial E_{kin}}{\partial x_i} \equiv m \ddot{x}_i = -\partial_x F(x),
\]

where \(\partial_s\) denotes the partial derivative with respect to the variable \(s\) (which is either space coordinate \(x\) or time \(t\)). This approach has been reformulated in Hamiltonian form for the purpose of development of the Human systems is sometimes called soliton theory. In this section, however, we are interested in solitons arising in electrical field theory (as travelling wave solutions of the classical field equations) and primarily in the role they play in quantization of electrical field theories, which is a different point of view from the one in classical soliton theory.
Quantum Leap

_Biodynamics Engine_, the leading human–motion simulator (developed by the authors at Defence Science & Technology Organisation, Australia).

### 5.4.2 Internal Myofibrillar Bio–Quantum Mechanics

At the same time, on the micro–level in the phase $EMS_{\text{paths}}$, instead of a single path defined by the Newtonian equation of motion for the whole muscle (5.13), we have an ensemble of fluctuating and crossing, fractal paths with weighted probabilities (of the unit total sum). This ensemble of micro–paths is defined by the simplest instance of our adaptive path integral (7.70), similar to the Feynman’s original sum over histories,

\[ CONTRACT_{\text{paths}} = \int D[wx] e^{iS[x]}, \quad (5.14) \]

where $D[wx]$ is a functional measure on the space of all weighted paths, and the exponential depends on the action $S[x]$ given by (7.45). In the language of transition–propagators, the integral over histories (7.47) can be decomposed into the product of myofibrillar action propagators. This procedure can be redefined in a mathematically cleaner way if we Wick–rotate the time variable $t$ to imaginary values $t \mapsto \tau = i t$, thereby making all integrals real:

\[ \int D[wx] e^{iS[x]} \xrightarrow{\text{Wick}} \int D[wx] e^{-S[x]}. \quad (5.15) \]

Discretization of (7.48) gives the standard thermodynamic partition function

\[ Z = \sum_j e^{-w_j E_j/T}, \quad (5.16) \]

where $E_j$ is the motion energy eigenvalue (reflecting each possible motivational energetic state), $T$ is the temperature environmental control parameter, and the sum runs over all motion energy eigenstates (labelled by the index $j$). From (5.16), we can further calculate all thermodynamical and statistical $EMS$–properties (see _Feynman (1972)_), as for example, _transition entropy $S = k_B \ln Z$, etc._

Now, both the action integral (7.45) and the path integral (7.47) are closely related to the _molecular soliton model of muscular contraction_, as

---

_Feynman propagators are otherwise called Fredholm kernels or Green’s functions._
described by the Korteweg–De Vries equation \cite{Ivancevic and Pearce (2001a)}

\[ f_t - 6f f_x + f_{xxx} = 0, \quad (f_x = \partial_x f), \]  

(5.17)

where \( x \in \mathbb{R} \) and \( f \) is a real–valued smooth function defined on \( \mathbb{R} \). It can also be described by nonlinear Schrödinger equation \cite{Ivancevic and Pearce (2001a)}

\[ i\partial_t \psi = -\partial_x^2 \psi + 2\chi |\psi|^2 \psi, \]  

(5.18)

for \(-\infty < x < +\infty\). Here \( \psi(x,t) \) is a smooth complex–valued wave function with initial condition \( \psi(x,t)|_{t=0} = \psi(x) \) and \( \chi \) is a nonlinear parameter. In the linear limit \( (\chi = 0) \) (5.18) becomes the ordinary Schrödinger equation for the wave function of the free 1D particle with mass \( m = 1/2 \).

It is clear that these two solitary equations have a quantum–mechanical origin\(^6\). Recall, that by the use of the first quantization method (instead of the Feynman integral), every classical biodynamic observable \( F \) is represented in the Hilbert space \( L^2(\psi) \) of square–integrable complex \( \psi \)–functions by a Hermitian (self–adjoint) linear operator \( \hat{F} \) with real eigenvalues. The classical Poisson bracket \( \{ F, G \} = K \) corresponds to the Dirac quantum commutator \( [ \hat{F}, \hat{G} ] = i\hat{K} \) (where, as always we have used normal units in which \( \hbar = 1 \) ). Therefore the classical Poisson evolution equation

\[ \dot{F} = \{ F, H \}. \]  

(5.19)

corresponds, in the Heisenberg picture, to the quantum evolution equation

\[ i\dot{\hat{F}} = [\hat{F}, \hat{H}], \]

for any representative operator \( \hat{F} \) and quantum Hamiltonian operator \( \hat{H} \).

By the Ehrenfest theorem (see, e.g., \cite{Feynman (1972)}) , this equation is also valid for expectation values \( \langle \cdot \rangle \) of observables, that is,

\[ i\langle \dot{\hat{F}} \rangle = \langle [\hat{F}, \hat{H}] \rangle. \]

For technical details on classical muscular mechanics, including the celebrated work of Nobel Laureates:

(1) the microscopic sliding filament model of A.F. Huxley;
(2) the macroscopic force–velocity model of A.V. Hill;

\(^6\)As Richard Feynman says in The Feynman Lectures on Physics: “Where did we get that [Schrödinger equation] from? It’s not possible to derive it from anything you know. It came out of the mind of Schrödinger.” Yet, Schrödinger equation can be (and usually is) derived from the Feynman path integral.
(3) the celebrated Hodgin–Huxley neural (and subsequently muscular) excitation model (A.L. Hodgin and A.F. Huxley); and
(4) the Eccles model of synaptic activation.

5.5 Anatomical Geometry of the Face & Body Shape: $EM^S_{geom}$

5.5.1 External Face & Body Geometry

On the macro–level in the phase $EM^S_{geom}$ representing a smooth skin 2D manifold–patch $M_2$ with the Riemannian metric tensor $g_{ij} = g_{ij}(M_2)$ defined at each local face or body point, we formulate the geometric action principle

$$\delta S[g_{ij}] = 0,$$

where $S = S[g_{ij}]$ is the 2D geodesic action on the surface $M$,

$$S[g_{ij}] = \int d^n x \sqrt{g_{ij} dx^i dx^j},$$

(Einstein’s summation convention over repeated indices is assumed). The corresponding Euler–Lagrangian equation gives the geodesic equation of the shortest path on the manifold $M_2$,

$$\ddot{x}^i + \Gamma^i_{jk} \dot{x}^j \dot{x}^k = 0,$$

where the symbol $\Gamma^i_{jk}$ denotes the so–called affine connection which is the source of curvature, which is geometric description for noise (see [Ingber (1997); Ingber (1998)]). The higher the local curvatures of the skin manifold–patch $M_2$, the greater the internal $EM^S$–noise. This noise is the source of our micro–level fat–related fluctuations.

Assuming that the electro–physiological principles of the $EM^S$–based body–shaping are identical (only less subtle) to the principles of the $EM^S$–based face–shaping, in the following subsections, we will focus on the facial anatomical geometry.

5.5.1.1 Facial Curvatures and Their Deviations

In this subsection we consider human face, with its distinguished local anatomical features, as a 2D Riemannian manifold (i.e, a smooth skin manifold–patch $M_2$), determined by muscular structural and functional
Table 5.1 Functional anatomy of the facial musculature (see any anatomical textbook.)

<table>
<thead>
<tr>
<th>Muscle</th>
<th>Origin</th>
<th>Insertion</th>
<th>Action</th>
<th>Innervation</th>
</tr>
</thead>
<tbody>
<tr>
<td>depressor anguli oris</td>
<td>oblique line of mandible</td>
<td>angle of mouth</td>
<td>pulls the corner of the mouth downward</td>
<td>marginal mandibular &amp; buccal branches of facial nerve (VII)</td>
</tr>
<tr>
<td>levator labii superioris</td>
<td>inferior margin of orbit</td>
<td>skin of upper lip</td>
<td>elevates upper lip</td>
<td>buccal branch of facial nerve (VII)</td>
</tr>
<tr>
<td>zygomaticus major</td>
<td>zygomatic bone, upper lateral surface</td>
<td>skin of angle of mouth</td>
<td>elevates and draws the corner of mouth laterally</td>
<td>zygomatic and buccal branches of facial nerve (VII)</td>
</tr>
<tr>
<td>orbicularis oris</td>
<td>skin and fascia of lips and area surrounding lips</td>
<td>skin and fascia of lips</td>
<td>purses the lips</td>
<td>buccal branch of facial nerve (VII)</td>
</tr>
<tr>
<td>buccinator</td>
<td>pterygomandibular raphe, mandible, and maxilla lateral to molar teeth</td>
<td>angle of mouth &amp; lateral portion of upper and lower lips</td>
<td>pulls corner of mouth laterally; pulls cheek against teeth</td>
<td>buccal branches of facial nerve (VII)</td>
</tr>
<tr>
<td>platysma</td>
<td>fascia overlying the pectoralis major and deltoid muscles</td>
<td>inferior border of mandible and skin of lower face</td>
<td>draws corners of mouth down; aids in depression of the mandible</td>
<td>facial nerve (VII), cervical branch</td>
</tr>
</tbody>
</table>

anatomy (see Figure 5.5). Here we demonstrate that this anatomical geometry is not static, but rather dynamic in an extremely complex way, which can be controlled by a proper EMS.

Recall from chapter 2, that Riemannian metric on any smooth nD Riemannian manifold $M$ is a positive–definite quadratic form $g : M \to \mathbb{R}$, which is in local coordinates $(x^1(s), \ldots, x^n(s))$, dependent on the affine line parameter $s$ at a point $m \in M$, defined as a symmetric $(0,2)$ tensor–field $g_{ij}(m) = g \left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right)(m)$, (see also [Boothby (1986); Ivancevic (2004); Ivancevic (2002); Ivancevic and Pearce (2001b)].

An infinitesimal distance between the two nearby local points $m$ and $n$
Fig. 5.5  Structural anatomy of the facial musculature (together with superficial branches of the facial nerve), showing local geometric features: distances (metrics), curvatures and directions for muscular contractions (see Table 5.1), modelled here as geodesic deviations. Here we show only facial EMS, assuming that all biophysical and geometric principles are the same (only less subtle) for body EMS.

on $M$ is defined by the line element

$$ds^2 = g_{ij} dx^i dx^j,$$

and realized by the geodesics $x^i(s)$. In local coordinates $(x^1(s),\ldots,x^n(s))$ at a point $m \in M$, the geodesic defining equation (5.21), derived from the geometric action principle (7.56), is a second order ordinary differential equation with the Christoffel symbols $\Gamma^i_{jk} = \Gamma^i_{jk}(m)$ of the affine (Levi–Civita) connection $\nabla_{x^i}$ are calculated at the point $m \in M$ with local coordinates $(x^1(s),\ldots,x^n(s))$.

If $\Gamma^i_{jk}(m) = 0$, the manifold $M$ is flat at the point $m$. This means that the Riemann curvature tensor, a symmetric $(1,3)$ tensor field $R^i_{jkl} = R^i_{jkl}(m)$, locally defined at a point $m \in M$ as

$$R^i_{jkl} = \partial_\mu \Gamma^i_{jk} - \partial_\mu \Gamma^i_{jl} + \Gamma^i_{km} \Gamma^m_{jl} - \Gamma^i_{km} \Gamma^m_{jk},$$
also vanishes at a point \( m \in M \), i.e., \( R^i_{jkl}(m) = 0 \).

Elliptic manifolds have positive curvature, i.e., \( R^i_{jkl}(m) > 0 \) and nearby geodesics are converging on it (see Figure 6.5), while hyperbolic ones have negative curvature, i.e., \( R^i_{jkl}(m) < 0 \) and nearby geodesics are diverging on it (see Figure 6.6). In case of a surface, which is a 2D manifold, the metric tensor is reduced to the scalar curvature \( R \). An example of elliptic (convex) surfaces is the sphere with the curvature \( R = \rho^{-2} \) (where \( \rho \) is the radius), while an example of hyperbolic (concave) surfaces is the Lobachevsky plane with the curvature \( R = -1 \).

Now, the rate of change of a vector field \( A^k \) on the manifold \( M \) along the curve \( x^i(s) \) is properly defined by the covariant derivative [Ivancevic and Ivancevic (2007b)]:

\[
\frac{D}{ds} A^k = \dot{x}^i \nabla_x A^i = \dot{x}^i \left( \partial_x A^k + \Gamma^k_{ij} A^j \right) = \dot{A}^k + \Gamma^k_{ij} \dot{x}^i A^j,
\]

which defines the parallel transport along the curve \( x^i(s) \) at a point \( m \in M \) as \( \frac{D}{ds} A^k(m) = 0 \).

By applying the previous result to itself, we can obtain an expression for the second covariant derivative of the vector field \( A^k \) along the curve \( x^i(s) \):

\[
\frac{D^2}{ds^2} A^k = \frac{d}{ds} \left( \dot{A}^k + \Gamma^k_{ij} \dot{x}^i A^j \right) + \Gamma^k_{ij} \dot{x}^i \left( \dot{A}^j + \Gamma^j_{mn} \dot{x}^m A^n \right).
\]

In the same local coordinates \( (x^1(s),...,x^n(s)) \) at a point \( m \in M \), let \( \delta x^i = \delta x^i(s) \) denote the vector–field of geodesic deviation, i.e., the infinitesimal vector–field describing both normal and tangential separation between the two neighboring geodesics, then the Jacobi equation of the geodesic deviation on the manifold \( M \) holds ([Arnold (1989)]):

\[
\frac{D^2}{ds^2} \delta x^i + R^i_{jkl} \dot{x}^j \delta x^k \dot{x}^l = 0.
\]

This equation describes the relative acceleration between two infinitesimally close facial geodesics, which is proportional both to the facial curvature (measured by the Riemann tensor \( R^i_{jkl} \) at a point \( m \in M \)), and to the geodesic deviation \( \delta x^i \). Solutions of the Jacobi equation (5.22) are called Jacobi fields, or Jacobi flows.
5.5.1.2 Facial Riemannian Patch–Manifold

Local anatomical features of the human face can be considered as a collection of local 2D Riemannian manifold–patches $M^i_2$, i.e., Riemannian patches, determined by skeletal and muscular anatomy. Each of these local patches $M^i_2$ is represented by its own local coordinates $(x^1(s), x^2(s))$, defined at a distinguished point $m$. We propose here a 2D Jacobi fields $\delta x^1, \delta x^2$ (see Figures 6.5 and 6.6) to model respectively normal and tangential components of contractions of the facial muscles.

In 2D, the Riemann curvature tensor simplifies into:

$$R^i_{jmn} = \frac{1}{2} R g^{ik} (g_{km} g_{jn} - g_{kn} g_{jm}),$$

where $R$ denotes the scalar curvature. Consequently the equation of geodesic deviation (5.22) also simplifies into

$$\frac{D^2}{ds^2} \delta x^i + \frac{R}{2} \delta x^i - \frac{R}{2} \dot{x}^j (g_{jk} \dot{x}^j \delta x^k) = 0. \quad (5.23)$$

Now, if we work in a local Cartesian coordinate system, defined at the tangent plane $T_m M_2$ at a point $m$ by an orthogonal projection imaging,
Fig. 5.7  Local coordinate chart defined in a neighborhood of a point \( m \) on the concave–hyperbolic patch-manifold \( M_2 \) of the facial musculature, together with the diverging geodesic deviation: its tangent component \( \delta x_1 \) and its normal component \( \delta x_2 \).

the covariant derivative \( \frac{D^2}{ds^2} \) reduces to the ordinary derivative \( \frac{d^2}{ds^2} \) (as the Christoffel symbols \( \Gamma^{i}_{jk} \) vanish) and the metric tensor \( g_{ij} \) reduces to identity matrix \( I_{ij} \), so our 2\( D \) equation of geodesic deviation (5.23) reduces into a simple second order ordinary differential equation in just two coordinates \( x^i \) (\( i = 1, 2 \))

\[
\frac{d^2}{ds^2} \delta x^i + \frac{R}{2} \delta x^i - \frac{R}{2} \dot{x}^j (I_{jk} \dot{x}^k \delta x^j) = 0.
\]

Also, if we require that the two nearby geodesics be nearly parallel, the last term in (5.23) vanishes, and we are left with

\[
\frac{D^2}{ds^2} \delta x^i + \frac{R}{2} \delta x^i = 0. \quad (5.24)
\]

Again, if we work in a locally Cartesian coordinate system, our flat 2\( D \) equation of geodesic deviation simplifies into harmonic oscillator in which the scalar curvature \( R/2 \) plays the role of the spring constant:

\[
\frac{d^2}{ds^2} \delta x^i + \frac{R}{2} \delta x^i = 0. \quad (5.25)
\]

Therefore, equations (5.24) and (5.23) could be respectively regarded
Quantum Leap

as the first–order and second–order perturbations of the linear oscillator equation (5.25). These three equations represent the three levels of detail in our modelling of the facial muscular movements. The oscillator equation (5.25) has a simple family of sinus functions (with certain amplitudes, frequencies and phases) as a solution, while the two nonlinear equations (5.24) and (5.23) could be numerically integrated for zero initial deviation and its velocity, using any explicit Runge–Kutta–like integrator (see, e.g. [Ivancevic and Snoswell (2001)]). Each of them describes the facial movement caused by muscular contraction dependent on its local curvature, i.e., anatomical shape. Also, all three geometric oscillators have kinetic and potential energies respectively defined as quadratic forms:

\[ E_{\text{kin}} = \frac{1}{2} g_{ij} \dot{x}^i \dot{x}^j, \quad U = \frac{1}{4} R g_{ij} \dot{x}^i \dot{x}^j, \]

and derived from the muscular action principle (7.45) above.

### 5.5.2 Cellular Muscle–Fat Geometry

On the micro–level in the phase \( EMS_{\text{geom}} \), we have an adaptive sum over fractal geometries, represented by the path integral over all regional Riemannian metrics \( g_{ij} = g_{ij}(x) \) varying from point to point inside an \( n \)-dimensional muscle–fat manifold \( M \), underlying the external skin surface,

\[ \text{SHAPE}_{\text{geom}} = \int \mathcal{D}[g_{ij}] e^{i S[g_{ij}]} \quad \text{Wick} \quad \int \mathcal{D}[w g_{ij}] e^{-S[g_{ij}]} , \quad (5.26) \]

where \( \mathcal{D}[g_{ij}] \) denotes diffeomorphism equivalence classes of metrics \( g_{ij}(x) \) of Skin.

To include the severe change of topological structure (e.g., a change in a number of holes) in the manifold \( M \), equation (7.58) can be extended as

\[ \text{SHAPE}_{\text{geom/top}} = \sum_{\text{topol.}} \int \mathcal{D}[w g_{ij}] e^{i S[g_{ij}]} , \quad (5.27) \]

where the topological sum is taken over all components of connectedness of the manifold \( M \) determined by its Euler characteristics [Ivancevic (2002)]. This type of integral defines the theory of fluctuating geometries, a propagator between \( (n-1) \)-dimensional boundaries of the \( n \)-dimensional manifold \( M \). One has to contribute a meaning to the integration over geometries. A key ingredient in doing so is to approximate in a natural way the smooth structures of the manifold \( M \) by piecewise linear structures (mostly using...
topological simplices $\Delta^7$. In this way, after the Wick–rotation (7.48), the integral (7.58–5.27) becomes a simple statistical system, given by partition function

$$Z = \sum_{\Delta} \frac{1}{C_{\Delta}} e^{-S_{\Delta}},$$

where the summation is over all triangulations $\Delta$ of the manifold $M$, while the number $C_{\Delta}$ is the order of the automorphism group of the performed triangulation.

---

7 This is called the simplicial approximation.
This page intentionally left blank
Chapter 6

Quantum Games and Quantum Computers

In this chapter we present modern quantum games and quantum computers.

6.1 Nash’s Classical Game Theory

The 1994 Nobel Prize for Economic Sciences was jointly awarded to John F. Nash, Jr. from Princeton, Reinhard Selten from Bonn and John C. Harsanyi from Berkeley. This Prize recognized the central importance of game theory in modern economic theory. Also, the timing of these awards had a historical significance, since that year was the fiftieth anniversary of the publication of the classical book [Neumann and Morgenstern (1944)] on cooperative game theory by John Von Neumann and Oskar Morgenstern.

The same situations that economists and mathematicians call games, psychologists usually call social situations. While game theory has applications to “games” such as poker and chess, the social situations are the core of modern research in game theory. Game theory has two main branches: Cooperative game theory focuses on the formation of coalitions and studies social situations axiomatically. Non–cooperative game theory models a social situation by specifying the options, incentives and information of the “players” and attempts to determine how they will play.

As the work of the other two 1994 Nobel Laureates was largely based on Nash’s non–cooperative game theory in general, and his governing equilibrium dynamics in particular, in this section we will review Nash’s work, discovering a way to predict the outcome of almost any type of strategic interaction.\footnote{J. Nash’s life motivated the Oscar–winning film “A Beautiful Mind”.

1523
6.1.1 Basics of Classical Game Theory

Since the work of John von Neumann, “games” have been a scientific metaphor for a much wider range of human interactions in which the outcomes depend on the interactive strategies of two or more persons, who have opposed or at best mixed motives. Among the issues discussed in game theory are the following:

1) What does it mean to choose actions “rationally” when outcomes depend on the actions chosen by others and when information is incomplete?

2) In “games” that allow mutual gain (or, mutual loss) is it “rational” to cooperate to realize the mutual gain (or avoid the mutual loss) or is it “rational” to act aggressively in seeking individual gain regardless of mutual gain or loss?

3) If the answers to 2) are “sometimes,” in what circumstances is aggression rational and in what circumstances is cooperation rational?

4) In particular, do ongoing relationships differ from one–off encounters in this connection?

5) Can moral rules of cooperation emerge spontaneously from the interactions of rational egoists?

6) How does real human behavior correspond to “rational” behavior in these cases?

7) If it differs, in what direction? Are people more cooperative than would be “rational?” More aggressive? Both?

Game theory starts from a description of the game. There are two distinct but related ways of describing a game mathematically. The extensive form is the most detailed way of describing a game. It describes play by means of a game tree that explicitly indicates when players move, which moves are available, and what they know about the moves of other players and nature when they move. Most important, it specifies the pay–offs that players receive at the end of the game [Levine (2001)].

Fundamental to game theory is the notion of a strategy. A strategy is a specification of how to play the game in every contingency, i.e., a set of instructions that a player could give to a friend or program on a computer so that the friend or computer could play the game on his behalf. Generally, strategies are contingent responses: in the game of chess, for example, a strategy should specify how to play for every possible arrangement of pieces on the board.

An alternative to the extensive form is the normal or strategic form – a description of a game by specifying the strategies and pay–offs. This is
Table 6.1 Tucker’s “Prisoner’s Dilemma” game.

<table>
<thead>
<tr>
<th>Player 1</th>
<th>Player 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>not confess</td>
<td>5,5</td>
</tr>
<tr>
<td>confess</td>
<td>9,0</td>
</tr>
<tr>
<td></td>
<td>0,9</td>
</tr>
<tr>
<td></td>
<td>1,1</td>
</tr>
</tbody>
</table>

less detailed than the extensive form, specifying only the list of strategies available to each player. Since the strategies specify how each player is to play in each circumstance, we can work out from the strategy profile specifying each player’s strategy what pay–off is received by each player. This map from strategy profiles to pay–offs is called the normal or strategic form. It is perhaps the most familiar form of a game, and is frequently given in the form of a game matrix. For example, in the celebrated Prisoner’s Dilemma game (Table 6.1) the two players are partners in a crime who have been captured by the police. Each suspect is placed in a separate cell, and offered the opportunity to confess to the crime. The rows of the matrix correspond to strategies of the first player. The columns are strategies of the second player. The numbers in the matrix are the pay–offs: the first number is the pay–off to the first player, the second the pay–off to the second player. Note that higher numbers are better (more utility); here, the total pay–off to both players is highest if neither confesses so each receives 5. However, game theory predicts that this will not be the outcome of the game (hence the dilemma). Each player reasons as follows: if the other player does not confess, it is best for me to confess (9 instead of 5). If the other player does confess, it is also best for me to confess (1 instead of 0). So no matter what I think the other player will do, it is best to confess. The theory predicts, therefore, that each player following his own self–interest will result in confessions by both players.

This game has fascinated game theorists for a variety of reasons. First, it is a simple representation of a variety of important situations. For example, instead of confess/not confess we could label the strategies “contribute to the common good” or “behave selfishly.” This captures a variety of situations economists describe as public goods prob-

2This is the way of describing a game by listing the players (or individuals) participating in the game, and for each player, listing the alternative choices (strategies) available to that player. In the case of a two–player game, the actions of the first player form the rows, and the actions of the second player the columns, of a matrix. The entries in the matrix are two numbers representing the utility or pay–off to the first and second player respectively.
lems. An example is the construction of a bridge. It is best for everyone if the bridge is built, but best for each individual if someone else builds the bridge. This is sometimes referred to in economics as an externality. Similarly, this game could describe the alternative of two firms competing in the same market, and instead of confess/not confess we could label the strategies “set a high price” and “set a low price.” Naturally, it is best for both firms if they both set high prices, but best for each individual firm to set a low price while the opposition sets a high price.

A second feature of this game is that it is self-evident how an intelligent individual should behave. No matter what a suspect believes his partner is going to do, it is always best to confess. If the partner in the other cell is not confessing, it is possible to get 9 instead of 5. If the partner in the other cell is confessing, it is possible to get 1 instead of 0. Yet the pursuit of individually sensible behavior results in each player getting only 1 unit of utility, much less than the 5 units each that they would get if neither confessed. This conflict between the pursuit of individual goals and the common good is at the heart of many game theoretic problems.

A third feature of this game is that it changes in a very significant way if the game is repeated, or if the players will interact with each other again in the future. Suppose for example that after this game is over, and the suspects either are freed or are released from jail they will commit another crime and the game will be played again. In this case in the first period the suspects may reason that they should not confess because if they do not their partner will not confess in the second game. Strictly speaking, this conclusion is not valid, since in the second game both suspects will confess no matter what happened in the first game. However, repetition opens up the possibility of being rewarded or punished in the future for current behavior, and game theorists have provided a number of theories to explain the obvious intuition that if the game is repeated often enough, the suspects ought to cooperate.

The previous example illustrates the central concept in game theory, that of an equilibrium. This is an example of a dominant strategy equilibrium, a strategy profile in which each player plays best-response that does not depend on the strategies of other players. Here, the incentive of each player to confess does not depend on how the other player plays. Dominant strategy is the most persuasive notion of equilibrium known to game theorists. In the experimental laboratory, however, players who play the prisoner’s dilemma sometimes cooperate. The view of game theorists is that this does not contradict the theory, so much as reflect the fact that players
in the laboratory have concerns besides monetary pay-offs. An important
current topic of research in game theory is the study of the relationship be-
tween monetary pay-offs and the utility pay-offs that reflect players’ real
incentive for making decisions.

As a contrast to the prisoner’s dilemma, consider the Battle of the Sexes
game (Table 6.2). The story goes that a husband and wife must agree
on how to spend the evening. The husband (player 1) prefers to go to
the ball–game (2 instead of 1), and the wife (player 2) to the opera (also
2 instead of 1). However, they prefer agreement to disagreement, so if
they disagree both get 0. This game does not admit a dominant strategy
equilibrium. If the husband thinks the wife’s strategy is to choose the opera,
his best response is to choose opera rather than ball–game (1 instead of 0).
Conversely, if he thinks the wife’s strategy is to choose the ball–game, his
best response is ball–game (2 instead of 0). While in the prisoner’s dilemma,
the best response does not depend on what the other player is thought to
be doing, in the battle of the sexes, the best response depends entirely on
what the other player is thought to be doing. This is sometime called a
coordination game to reflect the fact that each player wants to coordinate
with the other player.

For games without dominant strategies the equilibrium notion most
widely used by game theorists is that of Nash equilibrium, a strategy profile
in which each player plays a best–response to the strategies of other play-
er. The battle of the sexes game has two Nash equilibria: both go to the
opera, or both go to the ball game: if each expects the other to go to the
opera (ball–game) the best response is to go to the opera (ball–game). By
way of contrast, one going to the opera and one to the ball–game is not a
Nash equilibrium: since each correctly anticipates that the other is doing
the opposite, neither one is playing a best response.

Games with more than one equilibrium pose a dilemma for game the-
ory: how do we or the players know which equilibrium to choose? This

Table 6.2 “Battle of the Sexes” game.

<table>
<thead>
<tr>
<th>Player 1</th>
<th>Player 2</th>
<th>opera</th>
<th>ball–game</th>
</tr>
</thead>
<tbody>
<tr>
<td>opera</td>
<td>1,2</td>
<td>0,0</td>
<td></td>
</tr>
<tr>
<td>ball–game</td>
<td>0,0</td>
<td>2,1</td>
<td></td>
</tr>
</tbody>
</table>

3If there is a set of strategies with the property that no player can benefit by changing
her strategy while the other players keep their strategies unchanged, then that set of
strategies and the corresponding payoffs constitute the Nash equilibrium.
Table 6.3 “Matching Pennies,” or “Holmes–Moriarity” game.

<table>
<thead>
<tr>
<th></th>
<th>Player 1</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Canterbury</td>
<td>Paris</td>
</tr>
<tr>
<td>Player 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canterbury</td>
<td>–1,1</td>
<td>1,–1</td>
</tr>
<tr>
<td>Paris</td>
<td>1,–1</td>
<td>–1,1</td>
</tr>
</tbody>
</table>

question has been a focal point for research in game theory since its inception. Modern theorists incline to the view that equilibrium is arrived at through learning: people have many opportunities to play various games, and through experience learn which is the “right” equilibrium.

Consider the well-known Matching Pennies game (Table 6.3). There is, however, a more colourful story from Conan Doyle’s Sherlock Holmes story The Last Problem. Moriarity (player 2) is pursuing Holmes (player 1) by train in order to kill Holmes and save himself. The train stops at Canterbury on the way to Paris. If both stop at Canterbury, Moriarity catches Holmes and wins the game (–1 for Holmes, 1 for Moriarity). Similarly if both stop at Paris. Conversely, if they stop at different places, Holmes escapes (1 for Holmes and –1 for Moriarity). This is an example of a zero sum game: one player’s loss is another player’s gain. In the story, Holmes stops at Canterbury, while Moriarity continues on to Paris. But it is easy to see that this is not a Nash equilibrium: Moriarity should have anticipated that Holmes would get off at Canterbury, and so his best response was to get off also at Canterbury. As Holmes says “There are limits, you see, to our friend’s intelligence. It would have been a coup–de–maître had he deduced what I would deduce and acted accordingly.” However, this game does not have any Nash equilibrium: whichever player loses should anticipate losing, and so choose different strategy.

What do game theorists make of a game without a Nash equilibrium? The answer is that there are more ways to play the game than are represented in the matrix. Instead of simply choosing Canterbury or Paris, a player can flip a coin to decide what to do. This is an example of a random or mixed strategy, which simply means a particular way of choosing randomly among the different strategies. It is a mathematical fact, although not an easy one to prove, that every game with a finite number of players and finite number of strategies has at least one mixed strategy Nash equilibrium. The mixed strategy equilibrium of the matching pennies game is well

4If we add up the wins and losses in a game, treating losses as negatives, and we find that the sum is zero for each set of strategies chosen, then the game is a “zero-sum game.”
known: each player should randomize 50–50 between the two alternatives. If Moriarity randomizes 50–50 between Canterbury and Paris, then Holmes has a 50% chance of winning and 50% chance of losing regardless of whether he choose to stop at Canterbury or Paris. Since he is indifferent between the two choices, he does not mind flipping a coin to decide between the two, and so there is no better choice than for him to randomize 50–50 himself. Similarly when Holmes is randomizing 50–50, there is no better choice for Moriarity to do the same. Each player, correctly anticipating that his opponent will randomize 50–50 can do no better than to do the same. So perhaps Holmes (or Conan Doyle) is not such a clever game theorist after all [Levine (2001)].

Mixed strategy equilibrium points out an aspect of Nash equilibrium that is often confusing for beginners. Nash equilibrium does not require a positive reason for playing the equilibrium strategy. In matching pennies, Holmes and Moriarity are indifferent: they have no positive reason to randomize 50–50 rather than doing something else. However, it is only an equilibrium if they both happen to randomize 50–50. The central thing to keep in mind is that Nash equilibrium does not attempt to explain why players play the way they do. It merely proposes a way of playing so that no player would have an incentive to play differently. Like the issue of multiple equilibria, theories that provide a positive reason for players to be at equilibrium have been one of the staples of game theory research, and the notion of players learning over time has played a central role in this research.

6.1.2 Nash Equilibrium: The Focal Point of Game Theory

6.1.2.1 Background to Nash Equilibrium

To describe the spirit of the time, we might quote from R. Aumann [Aumann (1987)]: “The period of the late 40’s and early 50’s was a period of excitement in game theory. The discipline had broken out of its cocoon and was testing its wings. Giants walked the earth. At Princeton, John Nash laid the groundwork for the general non–cooperative theory and for cooperative bargaining theory. Lloyd Shapley defined the value for coalitional games, initiated the theory of stochastic games, co–invented the core with D.B. Gillies, and together with John Milnor developed the first game models with a continuum of players. Harold Kuhn reformulated the extensive form of a game, and worked on behavior strategies and perfect recall.
Al Tucker discovered the Prisoner’s Dilemma, and supported a number of young game theorists through the Office of Naval Research.\footnote{It is important to recognize that these results did not respond to some suggestion of von Neumann, nor did they follow work that he had outlined or proposed; rather they were revolutionary new ideas that ran counter to von Neumann’s theory \cite{Nobel Seminar (1994)}. In almost every instance, it was a repair of some inadequacy of the game theory as outlined by von Neumann and Morgenstern. All of the results cited in \cite{Aumann (1987)} were obtained by members of the Mathematics Department at Princeton University. At the same time, the RAND Corporation, funded by the US Air Force, which was to be for many years the other major center of game-theoretic research, had just opened its doors in Santa Monica.}

The twenty-year old John Nash came to Princeton in September of 1948. He came to the Mathematics Department with a one sentence letter of recommendation from R.L. Duffin of Carnegie Institute of Technology\footnote{Today, this is the Carnegie Mellon University.}. This letter said, simply: “This man is a genius.” The results for which he is being honored by Nobel Prize were obtained in his first fourteen months of graduate study \cite{Nobel Seminar (1994)}.

Although the speed with which Nash obtained these results is surprising, equally surprising and certainly less widely known is that Nash had already completed an important piece of work on bargaining while still an undergraduate at the Carnegie Institute of Technology. This work, a subsection for an elective course in international economics, possibly the only formal course in economics he has ever had, was done in complete ignorance of the work of von Neumann and Morgenstern. In short, when he did this work he didn’t know that game theory existed. This result, which is a model of theoretical elegance, posits \textit{four reasonable requirements} or axioms on any game solution \cite{Nash (1997)}:

\begin{enumerate}
\item Any solution should be invariant under positive linear affine transformations\footnote{A\textit{ffine transformation} is any transformation preserving collinearity (i.e., all points lying on a line initially still lie on a line after transformation) and ratios of distances (e.g., the midpoint of a line segment remains the midpoint after transformation). An affine transformation may also be thought of as a \textit{shearing transformation}.} of the utility functions;
\item The solution should be efficient in the sense of \textit{Pareto optimality}, i.e., the best that could be achieved without disadvantaging at least one group.\footnote{Pareto optimality is a measure of efficiency, named after Vilfredo Pareto. An outcome of a game is Pareto optimal if there is no other outcome that makes every player at least as well off and at least one player strictly better off. That is, a Pareto Optimal outcome cannot be improved upon without hurting at least one player. Often, a Nash Equilibrium is not Pareto Optimal implying that the players’ payoffs can all be increased.}
\end{enumerate}
(3) Irrelevant alternatives should not change the outcome of the solution; and

(4) Bargaining problems with symmetric outcome sets should have symmetric solutions.

If these four reasonable conditions are satisfied then there exists a unique solution to the game, namely, the outcome that maximizes the product of the players’ utilities.

The main result, the definition of a Nash equilibrium, and a proof of its existence had been completed prior to November 1949, the date of submission by Lefschetz to the National Academy of Sciences. The thesis [Nash (1950a)] itself was completed and submitted after the persistent urging and counsel of Professor Al Tucker. The formal rules at Princeton require that the thesis must be read by two professors, who prepare a report evaluating the work. In this case, the readers were Tucker and the statistician, John Tukey; the evaluation was written by Tucker himself. He wrote, “This is a highly original and important contribution to the Theory of Games. It develops notions and properties of “non–cooperative games,” finite n-person games which are very interesting in themselves and which may open up many hitherto untouched problems that lie beyond the zero-sum two-person case. Both in conception and in execution this thesis is entirely the author’s own” [Nobel Seminar (1994)].

The theory which fully occupies half of the von Neumann and Morgenstern book [Neumann and Morgenstern (1944)] deals with cooperative game theory envisaging coalitions, side-payments, and binding agreements. In addition, they proposed as a solution concept a notion we now call a “stable set”, which need not exist for every game. By contrast, Nash proved by page 6 of his thesis [Nash (1950a)] that every n-person finite non–cooperative game has at least one (Nash) equilibrium point. This is a profile of mixed strategies, one for each player, which is such that no player can improve his pay–off by changing his mixed strategy unilaterally [Nobel Seminar (1994)].

The Nash equilibrium is without doubt the single game theoretic solution concept that is most frequently applied in economics. Economic applications include oligopoly, entry and exit, market equilibrium, search, location, bargaining, product quality, auctions, insurance, principal–agent (problems), higher education, discrimination, public goods, what have you. On the political front, applications include voting, arms control and inspection, as well as most international political models (deterrence, etc.).
Biological applications all deal with forms of strategic equilibrium; they suggest an interpretation of equilibrium quite different from the usual overt rationalism [Aumann (1987)].

In the short period of 1950–53, John Nash published four brilliant papers [Nash (1950b); Nash (1950c); Nash (1951); Nash (1953); Nash (1997)], in which he made at least three fundamentally important contributions to game theory:

(1) He introduced the distinction between cooperative and non-cooperative games. The former are games in which the players can make enforceable agreements and can also make irrevocable threats to other players. That is to say, they can fully commit themselves to specific strategies. In contrast, in noncooperative games, such self-commitment is not possible.

(2) As a natural solution concept for non-cooperative games, he introduced the concept of equilibrium points ([Nash (1950b); Nash (1951)]), now usually described as Nash equilibria. He also established their existence in all finite games.

(3) As a solution concept for two-person cooperative games, he proposed the Nash bargaining solution, first for games with fixed threats [Nash (1950c)], and later also for games with variable threats [Nash (1953)]. He also showed that, in the latter case, the two players’ optimal strategies will have maximin and minimax properties.

Von Neumann and Morgenstern’s classical book contains an excellent mathematical analysis of one class of non-cooperative games, viz. of two-person zero-sum games and of the minimax solution for such games. It contains also an excellent mathematical discussion of one cooperative solution concept, that of stable sets, for many specific games. Yet, it so happens that the concept of two-person zero-sum games has very few real-life applications. The concept of stable sets has even fewer empirical applications.

Then appeared the so-called Nash’s program, which proposes [Nash (1951)]:

The writer has developed a “dynamical” approach to the study of cooperative games based on reduction to non-cooperative form. One proceeds by constructing a model of the pre-play negotiation so that the steps of (this) negotiation become moves in a larger non-cooperative game...describing the total situation. This larger game is then treated in terms of the theory of this subsection...and
if values are obtained (then) they are taken as the values of the cooperative game. Thus, the problem of analyzing a cooperative game becomes the problem of obtaining a suitable, and convincing, non-cooperative model for the negotiation.

When game theorists speak of ‘Nash’s program’, it is this two-paragraph passage they have in mind. That is to say, they are talking about the program of trying to reduce cooperative games to non-cooperative games by means of suitable non-cooperative models of the bargaining process among the players.

It is an interesting fact of intellectual history that Nash’s papers in the early 1950’s at first encouraged game theorists to cultivate cooperative and non-cooperative game theory as largely independent disciplines, with a concentration on cooperative theory. However, twenty-five years later they encouraged a shift to non-cooperative game theory and to non-cooperative models of the negotiations among the players.

One of Reinhard Selten’s important contributions was his distinction between perfect & imperfect Nash equilibria. It was based on his realization that even strategy combinations fully satisfying Nash’s definition of Nash equilibria might very well contain some irrational strategies. To exclude such imperfect Nash equilibria containing such irrational strategies, at first he proposed what now are called subgame-perfect equilibria.

A Nash equilibrium is defined as a strategy combination with the property that every player’s strategy is a best reply to the other players’ strategies. This of course is true also for Nash equilibria in mixed strategies. But in the latter case, besides his mixed equilibrium strategy, each player will also have infinitely many alternative strategies that are his best replies to the other players’ strategies. This will make such equilibria potentially unstable.

In view of this fact, it was desirable to show, that “almost all” Nash equilibria can be interpreted as strict equilibria in pure strategies of a suitably chosen game with randomly fluctuating pay-off functions.

When John Nash published his basic papers on ‘equilibrium points in n-person games’, and ‘non-cooperative games’, nobody would have foretold the great impact of Nash equilibrium on economics and social science in general. It was even less expected that Nash’s equilibrium point concept would ever have any significance for biological theory.

Originally, von Neumann and Morgenstern developed game theory as a mathematical method especially adapted to economics and social science in
general. In the introduction of their book, they emphasized their view that methods taken over from the natural sciences are inadequate for their purpose. They succeeded in creating a new method of mathematical analysis not borrowed from physics. In the case of game theory the flow of methodological innovation did not go in the usual direction from the natural to the social sciences but rather in the opposite one. The basis for this extremely successful transfer is the concept of Nash equilibrium.

In his Ph.D. dissertation [Nash (1950a)], John Nash provided two interpretations of his equilibrium concept for non-cooperative games, one rationalistic and one population-statistic:

- In the first, standard, rationalistic interpretation, one imagines that the game in question is played only once, that the participants are “rational,” and that they know the full structure of the game. However, Nash comments: “It is quite strongly a rationalistic and idealizing interpretation.”

- The second, population-statistic interpretation, which Nash calls the mass-action interpretation, was until recently largely unknown. Here Nash imagines that the game in question is played over and over again by participants who are not necessarily “rational” and who need not know the structure of the game:
  “It is unnecessary to assume that the participants have full knowledge of the total structure of the game, or the ability and inclination to go through any complex reasoning processes. But the participants are supposed to accumulate empirical information on the relative advantages of the various pure strategies at their disposal.

  To be more detailed, we assume that there is a population (in the sense of statistics) of participants for each position of the game. Let us also assume that the ‘average playing’ of the game involves $n$ participants selected at random from the $n$ populations, and that there is a stable average frequency with which each pure strategy is employed by the ‘average member’ of the appropriate population.

  Since there is to be no collaboration between individuals playing in different positions of the game, the probability that a particular $n$–tuple of pure strategies will be employed in a playing of the game should be the product of the probabilities indicating the chance of each of the $n$ pure strategies to be employed in a random playing’.”
The Rationalistic Interpretation

A non-cooperative game is given by a set of players, each having a set of strategies and a pay-off function. A strategy vector is a Nash equilibrium if each player’s strategy maximizes his pay-off if the strategies of the others are held fixed. In his Ph.D. thesis [Nash (1950a)], Nash introduces this concept and he derives several properties of it, the most important one being existence of at least one equilibrium for every finite game. In published work [Nash (1950b); Nash (1951)], Nash provides two alternative, elegant proofs, one based on Kakutani’s fixed point theorem, the other based directly on Brouwer’s fixed point theorem. These techniques have inspired many other existence proofs, for example, in the area of general equilibrium theory (see Debreu (1984)).

Nash’s motivation for his standard “rationalistic and idealizing interpretation” which is applicable to a game played just once, but which requires that the players are rational and know the full structure of the game, runs as follows:

“We proceed by investigating the question: What would be a “rational” prediction of the behavior to be expected of rational playing the game in question? By using the principles that a rational prediction should be unique, that the players should be able to deduce and make use of it, and that such knowledge on the part of each player of what to expect the others to do should not lead him to act out of conformity with the prediction, one is led to the concept of a solution defined before” [Nash (1950a)].

In other words, a theory of rational behavior has to prescribe the play of a Nash equilibrium since otherwise the theory is self-destroying. Note that the argument invokes three assumptions:

(i) players actively randomize in choosing their actions,
(ii) players know the game and the solution, and
(iii) the solution is unique.

Later work has scrutinized and clarified the role of each of these assumptions. Harsanyi [Harsanyi (1973)] showed that a mixed strategy of one player can be interpreted as the beliefs (conjectures) of the other players concerning his behavior. This reinterpretation provides a “Bayesian” foundation for mixed strategy equilibria and eliminates the intuitive difficulties associated with them.

Since the rationalistic justification of equilibria relies on uniqueness, multiplicity of equilibria is problematic. Nash remarks that it sometimes happens that good heuristic reasons can be found for narrowing down the
Table 6.4 Nash’s game.

<table>
<thead>
<tr>
<th>Player 1</th>
<th>Player 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>a</td>
<td>1,2</td>
</tr>
<tr>
<td>b</td>
<td>−4,−1</td>
</tr>
</tbody>
</table>

set of equilibria. One simple example that Nash provides [Nash (1951)] is given in Table 6.4. This game has equilibria at \((a, A)\) and \((b, B)\), as well as a mixed equilibrium. Nash writes that “empirical tests show a tendency toward \((a, A)\),” but he does not provide further details. One heuristic argument is that \((a, A)\) is less risky than \((b, B)\), an argument that is formalized by Harsanyi and Selten’s concept of risk dominance [Harsanyi and Selten (1988)]. It figures prominently both in the literature that builds on the “rationalistic interpretation” as well as in the literature that builds on the “mass–action” interpretation of Nash equilibrium.

6.1.2.3 The Mass–Action Interpretation

Notation and Preliminaries

Consider a finite \(n\)–player game \(G\) in normal (or strategic) form [Weibull (1994)]. Let \(A_i\) be the pure-strategy set of player position \(i \in I = \{1, \ldots, n\}\), \(S_i\) its mixed-strategy simplex, and \(S = \prod_{i \in I} S_i\), the polyhedron of mixed-strategy profiles. For any player position \(i\), pure strategy \(\alpha \in A_i\) and mixed strategy \(s_i \in S_i\), let \(s_{i\alpha}\) denote the probability assigned to \(\alpha\). A strategy profile \(s\) is called interior if all pure strategies are used with positive probability. The expected pay–off to player position \(i\) when a profile \(s \in S\) is played is denoted \(\pi_i(s)\), while \(\pi_{i\alpha}(s)\) denotes the pay–off to player \(i\) when he uses pure strategy \(\alpha \in A_i\) against the profile \(s \in S\). A strategy profile \(s \in S\) is a Nash equilibrium if and only if \(s_{i\alpha} > 0\) implies \(\pi_{i\alpha}(s) = \max_{\beta \in A_i} \pi_{i\beta}(s)\).

In the spirit of the mass–action interpretation, imagine that the game is played over and over again by individuals who are randomly drawn from (infinitely) large populations, one population for each player position \(i\) in the game. A population state is then formally identical with a mixed-strategy profile \(s \in S\), but now each component \(s_i \in S_i\) represents the distribution of pure strategies in player population \(i\), i.e., \(s_{i\alpha}\) is the probability that a randomly selected individual in population \(i\) will use pure strategy \(\alpha \in A\) (the so-called \(\alpha\)–strategist), when drawn to play the game.
In this interpretation $\pi_i(\alpha, s)$ is the (expected) pay–off to an individual in player population $i$ who uses pure strategy $\alpha$ and $\pi_i(s) \sum \beta s_i \pi_i(\beta, s)$ is the average (expected) pay–off in player population $i$, both quantities being evaluated in population state $s$.

Suppose that every now and then, say, according to a statistically independent Poisson process, each individual reviews her strategy choice. By the law of large numbers the aggregate process of strategy adaptation may then be approximated by deterministic flows, and these may be described in terms of ordinary differential equations.\footnote{Note that these conditions can be relaxed to stochastic flows, commonly described either by Langevin stochastic ODEs, or by corresponding Fokker–Planck PDEs (defining time evolution of the solution’s probability density).}

**Innovative Adaptation**

We first consider the case when strategy adaptation is memory-less in the sense that the time rate of strategy revision and the choice probabilities of strategy-reviewing individuals are functions of the current state $s$ (only) Weibull (1994):

$$\dot{s}_i(\alpha, t) = f_i(\alpha, s(t)),$$

(overdot denotes time derivative) for some functions $f_i : S \rightarrow \mathbb{R}$. The quantity $f_i(s)$ thus represents the net increase per time unit of the population share of $\alpha$–strategists in player population $i$ when the overall population state is $s$. The (composite) function $f$ is assumed to be Lipschitz continuous and such that all solution trajectories starting in $S$ remain forever in $S$. Such a function $f$ is usually called a vector–field for equation (6.1).

The class of population dynamics (6.1) clearly allows for an innovative element; some individuals may begin using earlier unused strategies, either intentionally, by calculation or experiment, or unintentionally, by mistakes or mutations. This apparently implies that only those population states that constitute Nash equilibria can be stationary. The requirement is simple: if there is some (used or unused) pure strategy that results in a pay–off above the current average pay–off in the player population in question, then some such pure strategy will grow in population share. Formally, for any population state $s \in S$ and player position $i \in I$, let $B_i(s)$ denote the (possibly empty) subset of better–than–average pure strategies,
Quantum Leap

\[ B_i(s) = \{ \alpha \in A_i : \pi_{\alpha}(s) > \pi_i(s) \} . \]

Inventiveness can then be formalized as

\[ [\text{IN}] : \text{If } B_i(s) \neq \emptyset, \text{ then } f_{\alpha}(s) > 0 \text{ for some } \alpha \in B_i(s). \]

This condition is, for instance, met if reviewing individuals move toward the best replies to the current population state. Note that [IN] requires no knowledge about pay-offs to other player positions, nor is any detailed knowledge of the pay-offs to one’s own strategy set necessary. It is sufficient that individuals on average tend to twitch toward some of the better-than-average performing strategies.

Proposition 1. Suppose \( f \) meets [IN]. If a population state \( s \) is stationary under the associated dynamics [6.1], then \( s \) constitutes a Nash equilibrium of \( G \).

In order to incorporate memory in the dynamic process of strategy adaptation, one may introduce real variables \( P_{i\alpha} \), one for each player position \( i \) and pure strategy \( \alpha \in A_i \), that represent the \( i \)th population’s recollection of earlier pay-offs to pure strategy \( \alpha \). Assume that the recalled pay-off to any pure strategy \( \alpha \) changes with time according to

\[ P_{i\alpha}(t) = h_{i\alpha}(\pi_{i\alpha}(s(t)), p_{i\alpha}(t), t), \]

where \( h_{i\alpha} \) is a Lipschitz continuous function such that the recalled pay-off changes only if the current pay-off differs from the recalled pay-off, i.e., if \( h_{i\alpha}(\pi_{i\alpha}, p_{i\alpha}, t) = 0 \) implies \( \pi_{i\alpha} = p_{i\alpha} \).

The full adaptation dynamics with memory is then a system of differential equations in the state vector \( x = (s, p) \), where \( p \) moves according to [6.2] and \( s \) moves according to

\[ \dot{s}_{i\alpha}(t) = f_{i\alpha}(s(t), p(t)). \]

A counterpart to the earlier requirement [IN] of inventiveness is: if there is some (used or unused) pure strategy which is recalled to result in a pay-off above the average of the currently recalled pay-offs in the player population in question, then some such pure strategy will increase its population share. Formally, for any state \( (s, p) \) and player position \( i \in I \), let \( B_i(s, p) = \{ \alpha \in A_i : p_{i\alpha} > \sum_{\beta \in A_i} s_{i\beta} p_{i\beta} \} \). Inventiveness can then be formalized as

\[ [\text{IN}'] : \text{If } B_i(s, p) \neq \emptyset, \text{ then } f_{i\alpha}(s, p) > 0 \text{ for some } \alpha \in B_i(s, p). \]

The following extension of Proposition 1 is straightforward:
Proposition 2. Suppose $f$ meets [IN']. If a population state $(s,p)$ is stationary under the associated dynamics (6.2,6.3), then $s$ constitutes a Nash equilibrium of $G$.

Imitative Adaptation

It may be argued that the above classes of population dynamics go somewhat beyond the spirit of the mass–action interpretation since they presume that individuals perform a certain amount of calculations. Therefore, now assume no memory and no inventiveness as defined above. Thus, individuals now switch only between strategies already in use, and they do so only on the basis of these strategies current performance. Technically, this means that the population dynamics (6.1) has a vector–field $f$ of the form [Weibull (1994)]

$$f_{i\alpha}(s) = g_{i\alpha}(s)s_{i\alpha}.$$ (6.4)

The involved functions $g_{i\alpha}$ are usually called growth–rate functions, $g_{i\alpha}(s)$ being the growth rate of the population share of pure strategy $\alpha$ in player population $i$ when the population state is $s$. No vector-field of the form (6.4) is innovative in the sense of condition [IN], because if all individuals in a player population initially use only one (or a few) pure strategy then they will continue doing so forever, irrespective of whether some unused strategy yields a high pay–off or not. Consequently, stationarity does not imply Nash equilibrium for the present class of dynamics, which is usually called imitative.

A prime example of such dynamics is the so-called replicator dynamics used in evolutionary biology [Taylor (1979)]. Here, pure strategies represent genetically programmed behaviors, reproduction is asexual, each offspring inherits its parents strategy, and pay–offs represent reproductive fitness. Thus $\pi_{i\alpha}(s)$ is the number of (surviving) offspring to an $\alpha$–strategist in population $i$, and $\pi_i(s)$ is the average number of (surviving) offspring per individual in the same population. In the standard version of the population model, each pure strategy’s growth rate is proportional to its current pay–off

$$g_{i\alpha}(s) = \pi_{i\alpha}(s) - \pi_i(s).$$ (6.5)

If there exists a pure strategy which results in a pay–off above average in its player population (whether this pure strategy be currently used or not), then some such pure strategy has a positive growth rate. Hence, if all such
strategies are present in the population, then some such population share will grow. Formally:

\[
\text{[POS]} : \text{If } B_i(s) \neq \emptyset, \text{ then } g_{\alpha}(s) > 0 \text{ for some } \alpha \in B_i(s).
\]

The next proposition establishes the following implications under pay-off positive imitation:

(a) If all strategies are present in a stationary population state, then this constitutes a Nash equilibrium;

(b) A dynamically stable population state constitutes a Nash equilibrium;

(c) If a dynamic solution trajectory starts from a population state in which all pure strategies are present and the trajectory converges over time, then the limit state is a Nash equilibrium.

Proposition 3. Suppose \( g \) meets [POS], and consider the associated population dynamics \( (6.1) \) where \( f \) is defined in \( (6.4) \).

(a) If \( s \) is interior and stationary, then \( s \) is a Nash equilibrium.

(b) If \( s \) is dynamically stable, then \( s \) is a Nash equilibrium.

(c) If \( s \) is the limit of some interior solution trajectory, then \( s \) is a Nash equilibrium.

This demonstrates that the mass–action interpretation of Nash equilibria is in stark contrast with the usual rationalistic interpretation, but is closely related to ideas in evolutionary game theory. It opens new avenues for equilibrium and stability analysis of social and economic processes, and suggests new ways to combine insights in the social and behavior sciences with economic theory \([\text{Weibull (1994)}]\).

**Nash Equilibrium in Human Performance**

We end this section on classical (Nash) game theory by giving an example of mixed strategies in the game of tennis.

Great sport players use a variety of moves to keep their opponents guessing. Game theory suggests that there is an optimal way to confuse one’s opponent. This can be illustrated with the game of tennis \([\text{Dixit and Nalebuff (1993)}]\).

Imagine that it is your serve. Your opponent’s forehand is stronger than their backhand. However, if you continually serve to their backhand then they will anticipate correctly and gain an advantage. What should you do? It turns out that you can increase your performance by systematically favoring one side, although in an unpredictable way.
Table 6.5: Probability that your opponent returns your serve.

<table>
<thead>
<tr>
<th>Opponent Anticipates</th>
<th>Your Aim for Serve</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>forehand</td>
</tr>
<tr>
<td>forehand</td>
<td>90%</td>
</tr>
<tr>
<td>backhand</td>
<td>30%</td>
</tr>
</tbody>
</table>

More specifically, consider the following opponent. Her forehand is her strong game. If she anticipates your serve correctly, her forehand return will be successful 90% of the time, while an anticipated backhand return will be successful only 60% of the time. Obviously, if she anticipates incorrectly she fares much worse. If she anticipates the backhand side and serve goes to the forehand side, she successfully returns 30% of the time, and if she anticipates the forehand side and serve is to the backhand side, she successfully returns 20% of the time.

Your opponent wants to maximize her success of return, while your goal is the exact opposite. Before the match, what strategy should you take? The chart (see Figure 6.1) gives you your best strategy. To the left of where the two lines intersect, your opponent does better if she anticipates backhand, to the right she does better if she anticipates forehand. The 40%:60% mixture is the only strategy the receiver cannot exploit to her advantage. Using this strategy, her success of return will be only 48%.

![Fig. 6.1 Best serve strategy.](image)

Both theoretical and empirical study was performed in Walker and Wooders (2001), where a similar game was used as a theoretical model of the tennis serve and its relation to the winning of points in a tennis
match. The assumption used was: Every point in a tennis match was played as a $2 \times 2$ constant-sum normal-form game with a unique equilibrium in strictly mixed strategies. The authors studied ten matches from the four “Grand Slam” tournaments and the ATP–Masters tournaments, using videotapes. They selected top matches, between players (Rosewall, Smith, Borg, McEnroe, Connors, Lendl, Edberg, Cash, Wilander, Becker, Sampras, Agassi, and Korda) who were well known to each other, which lasted long enough to have lots of points scored. They analyzed only first serves (if the first serve went out of bounds, there was a second serve, but it was not analyzed in this study).

The conclusions of the study are: The players did use the equilibrium mixing probabilities, but they tended to switch from one action to the other too often. Those two facts can be consistent because there are two ways to randomize with, for example, 50 percent probability between F and B: FBFBFBFB and FFFBBB. Game theory predicts that each pattern would be equally likely, since if one were more likely, the other player could use that fact to better predict what would happen. The authors found that the players used the FBFBFB switching style more often than game theory would predict.

6.2 Quantum Games

Recall from the previous section on classical game theory that the important two player game is the so-called Prisoners’ Dilemma game \cite{Tucker} \footnote{Recall that the story goes as follows: Alice and Bob are arrested for a minor crime. Despite lacking evidence, the prosecutor is convinced that they are guilty of a more severe crime and reveals the following three sets of strategies: If both defect, i.e. refuse to confess, they will equally be convicted for minor crimes. Alternatively, if both confess, i.e. cooperate, they will be convicted, albeit not charged with maximum penalty. The third option is for one of them to cooperate and for the other one to defect. If, e.g., Alice cooperates and Bob defects, Alice will be released shortly whereas Bob will have to face maximum penalty, and vice versa.}. In this game two players suspected of having committed a minor crime (in the following referred to as Alice and Bob) can independently decide whether they intend to “cooperate” or “defect”. Being well aware of the consequences of their decisions the players obtain a certain pay–off according to their respective choices. This pay–off provides a quantitative characterization of their personal preferences. Both players are assumed to want to maximize their individual pay–off, yet they must pick their choice while they have to take their choices without the other player’s decision.
The players face a dilemma since dilemma lies in the fact that rational reasoning in such a situation dictates the players to defect, although they would be much happier if they both cooperated.

Formally, Alice has two basic choices, meaning that she can select from two possible strategies $s_A = C$ (cooperation) and $s_A = D$ (defection). Bob may also take $s_B = C$ or $s_B = D$. The game is defined by these possible strategies on the one hand, and on the other hand by a specification of how to evaluate the pay-off once the combination of chosen strategies $(s_A, s_B)$ is known, i.e., the utility functions mapping $(s_A, s_B)$ on a number [Myerson (1991)]. The expected pay-off quantifies the preference of the players.

A (strategic-form) two player game $\Gamma = (\{A, B\}, S, U)$ is fully specified by the set $S$ of pairs of strategies $(s_A, s_B)$, the utility functions $U = (u_A, u_B)$, and additional rules of the game consistent with the set of strategies.$^{11}$

In this subsection, following [Eisert and Wilkens (2000)], we further develop the idea of identifying strategic moves with quantum operations as introduced by Eisert et al. (1999); Meyer and Schmidt (2000). This approach appears to be fruitful in at least two ways Eisert et al. (1999); Meyer and Schmidt (2000); Goldenberg et al. (1999). On the one hand several recently proposed applications of quantum information theory can already be conceived as competitive situations where several parties with more or less opposed motives interact. These parties may, for example, apply quantum operations on a bi-partite quantum system [Plenio and Vedral (1998b)]. In the same context, quantum cloning has been formulated as a game between two players Werner (1998). Similarly, eavesdropping in quantum cryptography Bennett and Brassard (1984) can be regarded as a game between the eavesdropper and the sender, and there are similarities of the extended form of quantum versions of games and quantum algorithms Shor (1997); Ekert (1991). On the other hand a generalization of the theory of decisions into the domain of quantum probabilities seems interesting, as the roots of game theory are partly in probability theory. In this context it is of interest to investigate what solutions are attainable if superpositions of strategies are allowed for Eisert et al. (1999); Meyer and Schmidt (2000).

Game theory does not explicitly concern itself with how the information is transmitted once a decision is taken. Yet, it should be clear that the practical implementation of any (classical) game inevitably makes use of

$^{11}$The set $\{A, B\}$ is the set of players.
the exchange of voting papers, faxes, emails, ballots, and the like. In the Prisoners’ Dilemma, e.g., the two parties have to communicate with an advocate by talking to her or by writing a short letter on which the decision is indicated. The process of decision making in a game clearly requires the transmission of information. However, game theory does not concern itself with how information is transmitted once a decision is taken.

In the Prisoners’ Dilemma the two parties clearly have to communicate with an advocate by talking to her or by writing a short letter on which the decision is indicated. While game theory (as an abstract theory which is broad in its scope) rarely bothers to deal with the transmission of information explicitly, it should be clear that the practical implementation of any (classical) game inevitably makes use of the exchange of voting papers, faxes, emails, ballots, and so on. That is, bearing in mind that a game is also about the transfer of information, it becomes legitimate to ask what happens if these carriers of information are taken to be quantum systems, quantum information being a fundamental notion of information.

By classical means a two player binary choice game may be played as follows: An arbiter takes two coins and forwards one coin each to the players. The players then receive their coin with head up and may keep it as it is (“cooperate”) or turn it upside down so that tails is up (“defection”). Both players then return the coins to the arbiter who calculates the players’ final pay–off corresponding to the combination of strategies he obtains from the players. Here, the coins serve as the physical carrier of information in the game. In a quantum version of such a game quantum systems would be used as such carriers of information. For a binary choice two player game an implementation making use of minimal resources involves two qubits as physical carriers.

6.2.1 Quantum Strategies

In this subsection, following [Meyer and Schmidt (2000)], we analyze the effectiveness of quantum strategies, exemplified in the simple game PQ – penny – flip: The starship Enterprise is facing some immanent—and apparently inescapable—calamity when Q appears on the bridge and offers to help, provided Captain Picard\footnote{Captain Picard and Q are characters in the popular American television (and movie) series Star Trek: The Next Generation whose initials and abilities are ideal for this illustration.} can beat him at penny flipping: Picard is to place a penny head up in a box, whereupon they will take turns
(Q, then Picard, then Q) flipping the penny (or not), without being able to see it. Q wins if the penny is head up when they open the box. This is a two–person zero–sum strategic game which might be analyzed traditionally using the payoff matrix:

\[
\begin{array}{cccc}
N & NF & FN & FF \\
N & -1 & 1 & 1 & -1 \\
F & 1 & -1 & -1 & 1 \\
\end{array}
\]

where the rows and columns are labelled by Picard’s and Q’s pure strategies, respectively; F denotes a flip and N denotes no flip; and the numbers in the matrix are Picard’s payoffs: 1 indicating a win and −1 a loss. For example, consider the top entry in the second column: Q’s strategy is to flip the penny on his first turn and then not flip it on his second, while Picard’s strategy is to not flip the penny on his turn. The result is that the state of the penny is, successively: H, T, T, T, so Picard wins.

Now it is natural to define a two dimensional vector space \( V \) with basis \( \{ H, T \} \) and to represent player strategies by sequences of \( 2 \times 2 \) matrices. That is, the matrices

\[
\begin{align*}
F &= H \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
N &= H \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}
\end{align*}
\]

correspond to flipping and not flipping the penny, respectively, since we define them to act by left multiplication on the vector representing the state of the penny [Meyer and Schmidt (2000)]. A mixed action is a convex linear combination of \( F \) and \( N \), which acts as a \( 2 \times 2 \) (doubly) stochastic matrix:

\[
\begin{pmatrix} 1 - p & p \\ p & 1 - p \end{pmatrix}
\]

if the player flips the penny with probability \( p \in [0, 1] \). A sequence of mixed actions puts the state of the penny into a convex linear combination \( aH + (1 - a)T \), \( 0 \leq a \leq 1 \), which means that if the box is opened the penny will be head up with probability \( a \).

However, Q is utilizing a quantum strategy, namely a sequence of unitary, rather than stochastic, matrices. In standard Dirac notation the basis of \( V \) is written \( \{|H\}, |T\} \). A pure quantum state for the penny is a linear
combination \(a |H\rangle + b |T\rangle\), \(a, b \in \mathbb{C}\), \(a \overline{a} + b \overline{b} = 1\), which means that if the box is opened, the penny will be head up with probability \(a \overline{a}\). Since the penny starts in state \(|H\rangle\), this is the state of the penny if Q’s first action is the unitary operation

\[
U_1 = U(a, b) = \begin{pmatrix} H & T \\ T & H \end{pmatrix} \begin{pmatrix} a & b \\ b & -\overline{a} \end{pmatrix}.
\]

On the other hand, Captain Picard is utilizing a classical probabilistic strategy in which he flips the penny with probability \(p\). After his action the penny is in a mixed quantum state, i.e., it is in the pure state \(b |H\rangle + a |T\rangle\) with probability \(p\) and in the pure state \(a |H\rangle + b |T\rangle\) with probability \(1 - p\). Mixed states are conveniently represented as density matrices, elements of \(V \otimes V^\dagger\) with trace 1; the diagonal entry \((i, i)\) is the probability that the system is observed to be in state \(|i\rangle\). The density matrix for a pure state \(|\psi\rangle \in V\) is the projection matrix \(|\psi\rangle \langle \psi|\) and the density matrix for a mixed state is the corresponding convex linear combination of pure density matrices. Unitary transformations act on density matrices by conjugation: The penny starts in the pure state \(\rho_0 = |H\rangle \langle H|\) and Q’s first action puts it into the pure state:

\[
\rho_1 = U_1 \rho_0 U_1^\dagger.
\]

Picard’s mixed action acts on this density matrix, not as a stochastic matrix on a probabilistic state, but as a convex linear combination of unitary (deterministic) transformations [Meyer and Schmidt (2000)]:

\[
\rho_2 = pF \rho_1 F^\dagger + (1 - p) N \rho_1 N^\dagger.
\]

For \(p = \frac{1}{2}\) the diagonal elements of \(\rho_2\) are each \(\frac{1}{2}\). If the game were to end here, Picard’s strategy would ensure him an expected payoff of 0, independently of Q’s strategy. In fact, if Q were to employ any strategy for which \(a \overline{a} \neq b \overline{b}\), Picard could obtain an expected payoff of \(|a \overline{a} - b \overline{b}| > 0\) by setting \(p = 0, 1\) according to whether \(b \overline{b} > a \overline{a}\), or the reverse. Similarly, if Picard were to choose \(p \neq \frac{1}{2}\), Q could obtain an expected payoff of \(2|p - 1|\) by setting \(a = 1\) or \(b = 1\) according to whether \(p < \frac{1}{2}\), or the reverse. Thus the mixed/quantum equilibria for the two–move game are pairs \((|\frac{1}{2} F + \frac{1}{2} N|, [U(a, b)])\) for which \(a \overline{a} = b \overline{b}\) and the outcome is the same as if both players utilize optimal mixed strategies.
But Q has another move $U_3$ which again transforms the state of the penny by conjugation to $\rho_3 = U_3\rho_2 U_3^\dagger$. If Q’s strategy consists of $U_1 = U(1/\sqrt{2}, 1/\sqrt{2}) = U_3$, his first action puts the penny into a simultaneous eigenvalue 1 eigenstate of both $F$ and $N$, which is therefore invariant under any mixed strategy $pF + (1 - p)N$ of Picard; and his second action inverts his first to give $\rho_3 = |H\rangle\langle H|$. That is, with probability 1 the penny is head up! Since Q can do no better than to win with probability 1, this is an optimal quantum strategy for him. All the pairs $(pF + (1 - p)N, [U(1/\sqrt{2}, 1/\sqrt{2}), U(1/\sqrt{2}, 1/\sqrt{2})])$ are mixed/quantum equilibria for $PQ$ – penny – flip, with value $-1$ to Picard; this is why he loses every game.

Therefore, $PQ$ – penny – flip is a very simple game, but it is structurally similar to the oracle problems for which efficient quantum algorithms are known—with Picard playing the role of the oracle. In Simon’s problem the functions $f : \{0,1\}^n \rightarrow \{0,1\}^n$ which satisfy $f(x) = f(y)$ if and only if $y = x \oplus s$ for some $s \in \{0,1\}^n$ ($\oplus$ denotes componentwise addition, mod 2), correspond to Picard’s pure strategies; we may imagine the oracle choosing a mixed strategy intended to minimize our chances of efficiently determining $s$ probabilistically. Simon’s algorithm is a quantum strategy which is more successful than any mixed, i.e., probabilistic, one. Similarly, in the problem of searching a database of size $N$, the locations in the database correspond to pure strategies; again we may imagine the oracle choosing a mixed strategy designed to frustrate our search for an item at some specified location. Grover’s algorithm is a quantum strategy for a game of $2m$ moves alternating between us and the oracle, where $m = O(\sqrt{N})$, which outperforms any mixed strategy. These three examples suggest the following theorem: There is always a mixed/quantum equilibrium for a two–person zero–sum game, at which the expected payoff for the player utilizing a quantum strategy is at least as great as his expected payoff with an optimal mixed strategy. For the proof, see Meyer and Schmidt (2000).

Another natural question to ask is what happens if both players utilize quantum strategies. Meyer proposes the following statement: A two–person zero–sum game need not have a quantum/quantum equilibrium. That is, the situation when both players utilize quantum strategies is the same as when they both utilize pure (classical) strategies: there need not be any equilibrium solution. This suggests looking for the analogue of von Neumann’s result on the existence of mixed strategy equilibria. So we should consider strategies which are convex linear combinations of unitary actions—mixed quantum strategies: A two–person zero–sum game always has a
mixed quantum/mixed quantum equilibrium. For the proofs, see Meyer and Schmidt (2000).

6.2.2 Quantum Games

A quantum game could be any quantum system where more than one party can implement quantum operations and where the utility of the moves can be quantified in an appropriate manner.

A two–player quantum game $\Gamma = (\mathcal{H}, \rho, S_A, S_B, P_A, P_B)$ is completely specified by the underlying Hilbert space $\mathcal{H}$ of the physical system, the initial state $\rho \in S(\mathcal{H})$, where $S(\mathcal{H})$ is the associated state space, the sets $S_A$ and $S_B$ of permissible quantum operations of the two players, and the utility functionals $P_A$ and $P_B$, which specify the expected pay–off utility for each player. A quantum strategy $s_A \in S_A, s_B \in S_B$ is a quantum operation, that is, a completely positive trace–preserving map mapping the state space on itself. The quantum game’s definition also includes certain implicit rules, such as the order of the implementation of the respective quantum strategies. Rules also exclude certain actions, as the alteration of the pay–off during the game.

The quantum games proposed in Eisert et al. (1999) Meyer and Schmidt (2000) Goldenberg et al. (1999) can be cast into this form. Also, the quantum cloning device as described in Werner (1998) can be said to be a quantum game in this sense and additional rules of the game consistent with the set of quantum strategies. The utility functionals $P_A, P_B \to \mathbb{R}$ map the final state after applying the respective strategies on the set of real numbers; they quantify the individual preference of each player. The rules of the game include the order in which the two players may implement their respective quantum strategy. Rules also exclude certain actions, as the alteration of the pay–off during the game. A quantum game is called zero–sum game, if the expected pay–offs sum up to zero for all pairs of strategies, that is, if $P_A(s_A, s_B) = -P_B(s_A, s_B)$ for all $s_A \in S_A, s_B \in S_B$. Otherwise, it is called a non–zero sum game.

Note that we do not require a set of allowed strategies for a player to be a closed set. The identity operation is included, as “not to do anything” always seems to be a natural option, but in the following it is not assumed.

---

$^{13}$E.g., Alice may allow a coupling of the original quantum system to an auxiliary quantum system and let the two unitarily interact. After performing a projective measurement on the composite system she could eventually consider the original system again by taking a partial trace with respect to the auxiliary part.
that these sets are closed sets. The allowed set of operations does not necessarily include only unitary operations, but any trace–preserving completely positive map corresponding with a generalized measurement.

It is natural to call two quantum strategies of Alice $s_A$ and $s'_A$ equivalent, if $P_A(s_A, s_B) = P_A(s'_A, s_B)$ and $P_B(s_A, s_B) = P_A(s'_A, s_B)$ for all possible $s_B$. That is, if $s_A$ and $s'_A$ yield the same expected pay–off for both players for all allowed strategies of Bob. In the same way strategies $s_B$ and $s'_B$ of Bob will be identified.

A solution concept provides advice to the players with respect to the action they should take. Such a solution must be based on the particular utility functionals which define the preference of each player. Indeed, in the Prisoners’ Dilemma game it was obvious what to accept as a solution, as to “defect” implied an advantage, regardless of the strategy of the other player. These definitions are fully analogous to the corresponding definitions in standard game theory [Myerson (1991)].

A quantum strategy $s_A$ (implemented by Alice) is called dominant strategy of Alice if

$$P_A(s_A, s'_B) \geq P_A(s'_A, s'_B)$$

for all $s'_A \in S_A$, $s'_B \in S_B$. Analogously we can define a dominant strategy for Bob. A pair $(s_A, s_B)$ is said to be an equilibrium in dominant strategies if $s_A$ and $s_B$ are the players’ respective dominant strategies. A combination of strategies $(s_A, s_B)$ is called a Nash equilibrium if

$$P_A(s_A, s_B) \geq P_A(s'_A, s_B),$$
$$P_B(s_A, s_B) \geq P_B(s_A, s'_B)$$
for all $s'_A \in S_A$, $s'_B \in S_B$. $(s_A, s_B)$ is a solution in maxi-min-strategies, if

$$\min_{s'_B \in S_B} P_A(s_A, s'_B) \geq \min_{s'_B \in S_B} P_A(s'_A, s'_B) \quad \text{for all } s'_A \in S_A,$$

$$\min_{s'_A \in S_A} P_B(s'_A, s_B) \geq \min_{s'_A \in S_A} P_B(s'_A, s'_B) \quad \text{for all } s'_B \in S_B.$$ 

A pair of strategies $(s_A, s_B)$ is called Pareto optimal, if it is not possible to increase one player’s pay–off without lessening the pay–off of the other player. A solution in dominant strategies is the strongest solution concept for a non–zero sum game. In the Prisoner’s Dilemma defection is the dominant strategy, as it is favorable regardless what strategy the other party picks. A dominant strategy exists if a particular strategy provides the maximum pay–off regardless of the strategy selected by the other player.

Typically, however, the optimal strategy depends on the strategy chosen by the other party. A Nash equilibrium implies that neither player has a motivation to unilaterally alter his or her strategy from this equilibrium solution, as this action will lessen his or her pay–off. Given that the other player will stick to the strategy corresponding to the equilibrium, the best result is achieved by also playing the equilibrium solution. The concept of Nash equilibria is of paramount importance to studies of non–zero–sum games. It is, however, only an acceptable solution concept if the Nash equilibrium is unique. For games with multiple equilibria the application of a hierarchy of natural refinement concepts may finally eliminate all but one of the Nash equilibria. Note that a Nash equilibrium is not necessarily efficient. In the Prisoners’ Dilemma for example there is a unique equilibrium, but it is not Pareto optimal, meaning that there is another outcome which would make both players better off.

### 6.2.3 Two–Qubit Quantum Games

In the subsequent investigation we turn to specific games where the classical version of the game is faithfully entailed in the quantum game. In a quantum version of a binary choice game two qubits are prepared by a arbiter in a particular initial state, the qubits are sent to the two players who have physical instruments at hand to manipulate their qubits in an appropriate manner. Finally, the qubits are sent back to the arbiter who performs a measurement to evaluate the pay–off. The source, the devices to manipulate the state, and the measurement apparatus are known to the players when they are obliged to take their decision.

For such a bi–partite quantum game the system of interest is a quantum
system with underlying Hilbert space
\[ \mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \]
and associated state space \( S(\mathcal{H}) \). Quantum strategies \( s_A \) and \( s_B \) of Alice and Bob are local quantum operations acting in \( \mathcal{H}_A \) and \( \mathcal{H}_B \) respectively.\footnote{The quantum strategies \( s_A \otimes 1 \) and \( 1 \otimes s_B \) are in the following identified with \( s_A \) and \( s_B \), respectively.}

That is, they are completely positive maps of the form
\[ s_A \otimes s_B : S(\mathcal{H}) \to S(\mathcal{H}). \]

That is, Alice and Bob are restricted to implement their respective quantum strategy \( s_A \) and \( s_B \) on their qubit only. In this step they may choose any quantum strategy that is included in the set of strategies \( S \). They are both well aware of the set \( S \), but they do not know which particular quantum strategy the other party would actually implement. As the application of both quantum strategies amounts to a map \( s_A \otimes s_B : S(\mathcal{H}) \to S(\mathcal{H}) \), after execution of the moves the system is in the state
\[ \sigma = (s_A \otimes s_B)(\rho). \]

Particularly important will be unitary operations \( s_A \) and \( s_B \). They are associated with unitary operators \( U_A \) and \( U_B \), written as \( s_A \sim U_A \) and \( s_B \sim U_B \). In this case the final state \( \sigma \) is given by
\[ \sigma = (U_A \otimes U_B)\rho(U_A \otimes U_B)^\dagger. \]

From now on both the sets of strategies of Alice and Bob and the pay–off functionals are taken to be identical, that is,
\[ S_A = S_B = S \quad \text{and} \quad P_A = P_B = P, \]
such that both parties face the same situation.

The quantum game \( \Gamma = (\mathbb{C}^2 \otimes \mathbb{C}^2, \rho, S, P, P) \) can be played in the following way: The initial state \( \rho \) is taken to be a maximally entangled state in the respective state space. In order to be consistent with Eisert \emph{et al.} (1999) let \( \rho = |\psi\rangle \langle \psi| \) with
\[ |\psi\rangle = (|00\rangle + i|11\rangle)/\sqrt{2}, \]
where the first entry refers to \( \mathcal{H}_A \) and the second to \( \mathcal{H}_B \). The expected pay–off is determined from this state \( \sigma \). The two qubits are forwarded to the arbiter who performs a projective selective measurement on the final
state $\sigma$ with the so-called Kraus operators, $\pi_{CC}$, $\pi_{CD}$, $\pi_{DC}$, and $\pi_{DD}$, where

$$
\begin{align*}
\pi_{CC} &= |\psi_{CC}\rangle \langle \psi_{CC}|, \quad |\psi_{CC}\rangle = (|00\rangle + i|11\rangle)/\sqrt{2}, \\
\pi_{CD} &= |\psi_{CD}\rangle \langle \psi_{CD}|, \quad |\psi_{CD}\rangle = (|01\rangle - i|10\rangle)/\sqrt{2}, \\
\pi_{DC} &= |\psi_{DC}\rangle \langle \psi_{DC}|, \quad |\psi_{DC}\rangle = (|10\rangle - i|01\rangle)/\sqrt{2}, \\
\pi_{DD} &= |\psi_{DD}\rangle \langle \psi_{DD}|, \quad |\psi_{DD}\rangle = (|11\rangle + i|00\rangle)/\sqrt{2}.
\end{align*}
$$

According to the outcome of the measurement, a pay-off of $A_{CC}$, $A_{CD}$, $A_{DC}$, or $A_{DD}$ is given to Alice, Bob receives $B_{CC}$, $B_{CD}$, $B_{DC}$, or $B_{DD}$. The utility functionals, also referred to as expected pay-off of Alice and Bob, read

$$
\begin{align*}
P_A(s_A, s_B) &= A_{CC} \text{tr}[\pi_{CC}\sigma] + A_{CD} \text{tr}[\pi_{CD}\sigma] + A_{DC} \text{tr}[\pi_{DC}\sigma] + A_{DD} \text{tr}[\pi_{DD}\sigma], \\
P_B(s_A, s_B) &= B_{CC} \text{tr}[\pi_{CC}\sigma] + B_{CD} \text{tr}[\pi_{CD}\sigma] + B_{DC} \text{tr}[\pi_{DC}\sigma] + B_{DD} \text{tr}[\pi_{DD}\sigma].
\end{align*}
$$

It is important to note that the Kraus operators are chosen in such a way that the classical game is fully entailed in the quantum game: The classical strategies $C$ and $D$ are associated with particular unitary operators with a matrix representation

$$
C \sim \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad D \sim \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.
$$

$C$ does not change the state at all, $D$ implements a “spin–flip”. If both parties stick to these classical strategies, equations (6.8) and (6.9) guarantee that the expected pay-off is exactly the pay-off of the corresponding classical game defined by the numbers $A_{CC}$, $A_{CD}$, $A_{DC}$, $A_{DD}$, $B_{CC}$, $B_{CD}$, $B_{DC}$, and $B_{DD}$. E.g., if Alice plays $C$ and Bob chooses $D$, the state $\sigma$ after implementation of the strategies is given by

$$
\sigma = (C \otimes D)(\rho) = |\psi_{CD}\rangle \langle \psi_{CD}|,
$$

such that Alice obtains $A_{CD}$ units and Bob $B_{CD}$ units pay-off. In this way the peculiarities of strategic moves in the quantum domain can be adequately studied. The players may make use of additional degrees of freedom which are not available with randomisation of the classical strategies, but they can also stick to mere classical strategies. This scheme can
be applied to any two player binary choice game and is to a high extent canonical.

### 6.2.3.1 Prisoners’ Dilemma

We now investigate the solution concepts for the quantum analogue of the Prisoners’ Dilemma as mentioned above\textsuperscript{15}

\[
A_{CC} = B_{CC} = 3, \quad A_{DD} = B_{DD} = 1, \\
A_{CD} = B_{DC} = 0, \quad A_{DC} = B_{CD} = 5.
\]

The solution provided by the introduced concepts clearly depends on the rules of the game, and in particular, on the sets of strategies. In all of the following sets of allowed strategies $S$ the classical options (to defect and to cooperate) are included. Several interesting sets of strategies and concomitant solution concepts will at this point be studied. The first three subsections involve local unitary operations only, while in the last subsection other quantum operations are considered as well.

#### One–parameter set of strategies

The first set of strategies $S^{(CL)}$ involves quantum operations $s_A$ and $s_B$ which are local rotations with one parameter. The matrix representation of the corresponding unitary operators is taken to be

\[
U(\theta) = \begin{pmatrix}
\cos(\theta/2) & \sin(\theta/2) \\
-\sin(\theta/2) & \cos(\theta/2)
\end{pmatrix},
\]

with $\theta \in [0, \pi]$. Hence, in this simple case, selecting strategies $s_A$ and $s_B$ amounts to choosing two angles $\theta_A$ and $\theta_B$. The classical strategies of defection and cooperation are also included in the set of possible strategies. We associate cooperation with the operation $C$ and defection with $D$, defined as

\[
C \sim U(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad D \sim U(\pi) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.
\]

\textsuperscript{15}From a game theoretical viewpoint, any positive numbers satisfying the symmetry conditions $A_{CC} = B_{CC}, A_{DD} = B_{DD}, A_{CD} = B_{DC}, A_{DC} = B_{CD}$ and the inequalities $A_{DC} > A_{CC} > A_{DD} > A_{CD}$ and $A_{CC} \geq (A_{CD} + A_{DC})/2$ define a (strong) Prisoners’ Dilemma.
Quantum Leap

An analysis of the expected pay-offs $P_A$ and $P_B$,

$$P_A(\theta_A, \theta_B) = 3|\cos(\theta_A/2) \cos(\theta_B/2)|^2 + 5|\cos(\theta_B/2) \sin(\theta_A/2)|^2,$$

$$P_B(\theta_A, \theta_B) = 3|\cos(\theta_A/2) \cos(\theta_B/2)|^2 + 5|\sin(\theta_B/2) \cos(\theta_A/2)|^2 + |\sin(\theta_A/2) \sin(\theta_B/2)|^2,$$

shows that this game is the classical Prisoners’ Dilemma game [Eisert et al. (1999)]. The pay-off functions are actually identical to the analogous functions in the ordinary Prisoners’ Dilemma with mixed (randomised) strategies, where cooperation is chosen with the classical probability $p = \cos^2(\theta/2)$. The inequalities

$$P_A(D, s_B) \geq P_A(s_A, s_B), \quad P_B(s_A, D) \geq P_B(s_A, s_B),$$

hold for all $s_A, s_B \in S^{(CL)}$, therefore, $(D, D)$ is an equilibrium in dominant strategies and thus the unique Nash equilibrium. As explained in the introduction this equilibrium is far from being efficient, because $P_A(D, D) = P_B(D, D) = 1$ instead of the Pareto optimal pay-off which would be 3.

Two-parameter set of strategies

A more general set of strategies is the following two-parameter set $S^{(TP)}$. The matrix representation of operators corresponding to quantum strategies from this set is given by An even more efficient solution is achieved when one considers the following two-parameter set $S^{(TP)}$ of local unitary operations. The matrix representation of the operators corresponding to quantum strategies $s_A$ and $s_B$ is given by Indeed, in this case the Pareto optimum can be realised. Both $s_A$ and $s_B$ are taken from a certain subset of local unitary operations. The matrix representation of the operators forming a subset of $SU(2)$ is given by

$$U(\theta, \phi) = \begin{pmatrix} e^{i\phi} \cos(\theta/2) & \sin(\theta/2) \\ -\sin(\theta/2) & e^{-i\phi} \cos(\theta/2) \end{pmatrix},$$

with $\theta \in [0, \pi]$ and $\phi \in [0, \pi/2]$. Selecting a strategy $s_A, s_B$ then means choosing appropriate angles $\theta_A, \phi_A$ and $\theta_B, \phi_B$. The classical pure strategies can be realised as

$$C \sim U(0, 0) \quad \text{and} \quad D \sim U(\pi, 0).$$
This case has also been considered in Eisert et al. (1999). The expected pay–off for Alice, e.g., explicitly reads
\[
P_A(\theta_A, \phi_A, \theta_B, \phi_B) = 3 |\cos(\phi_A + \phi_B) \cos(\theta_A/2)\cos(\theta_B/2)|^2 \\
+ 5 |\sin(\phi_A) \cos(\theta_A/2) \sin(\theta_B/2) - \cos(\phi_B) \cos(\theta_B/2) \sin(\theta_A/2)|^2 \\
+ |\sin(\phi_A + \phi_B) \cos(\theta_A/2) \cos(\theta_B/2) + \sin(\theta_A/2) \sin(\theta_B/2)|^2.
\] (6.10)

Due to entanglement of the initial state as defined in the general setup (6.6) Alice’s expected pay–off depends also on Bob’s choice. It turns out that the previous Nash equilibrium \((D, D)\) of \(S(\text{CL})\) is no longer an equilibrium solution, as both players can benefit from deviating from \(D\). However, concomitant with the disappearance of this solution another Nash equilibrium has emerged, given by \((Q, Q)\). The strategy \(Q\) is associated with a matrix \(Q = U(0, \pi/2) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\).

This Nash equilibrium is unique [Eisert et al. (1999)] and serves as the only acceptable solution of the game. The astonishing fact is that \(P_A(Q, Q) = P_B(Q, Q) = 3\) (instead of 1) so that the Pareto optimum is realised. No player could gain without lessening the other player’s expected pay–off. In this sense one can say that the dilemma of the original game has fully disappeared. In the classical game only mutual cooperation is Pareto optimal, but this pair of strategies does not correspond to a Nash equilibrium.

**General unitary operations**

One can generalise the previous setting to the case where Alice and Bob can implement operations \(s_A\) and \(s_B\) taken from \(S(\text{GU})\), where \(S(\text{GU})\) is the set of general local unitary operations, \(\mathcal{H}_A\) and \(\mathcal{H}_B\), respectively. Here, it could be suspected that the solution becomes more efficient the larger the sets of allowed operations are. But this is not the case. The previous Pareto optimal unique Nash equilibrium \((Q, Q)\) ceases to be an equilibrium solution if the set is enlarged: For any strategy \(s_B \in S(\text{GU})\) there exists an optimal answer local unitary operation \(s_A \in S(\text{GU})\) of Alice resulting in
\[
(s_A \otimes s_B)(\rho) = |\psi_{DC}\rangle\langle\psi_{DC}|,
\]
(6.11)
with \(\rho\) given in (6.6). This quantum strategy is given by and \(|\psi_{CD}\rangle\) in (6.7): For \(s_A\) That is, for any strategy of Bob \(s_B\) there is a strategy \(s_A\) of
Alice such that
\[ P_A(s_A, s_B) = 5 \quad \text{and} \quad P_B(s_A, s_B) = 0. \]

Take
\[ s_A \sim \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad s_B \sim \begin{pmatrix} -ib & a \\ -d & -ic \end{pmatrix}, \]
where \( a, b, c, d \) are appropriate complex numbers. Given that Bob plays the strategy \( s_B \) associated with a particular Nash equilibrium \( (s_A, s_B) \), Alice can always apply the optimal answer \( s_A \) to achieve the maximal possible pay-off. However, the resulting pair of quantum strategies can not be an equilibrium since again, the game being symmetric, Bob can improve his pay-off by changing his strategy to his optimal answer \( s'_B \). Hence, there is no pair \( (s_A, s_B) \) of pure strategies with the property that the players can only lose from unilaterally deviating from this pair of strategies. In particular, one can always gain from choosing a different strategy than the pure strategy of other party.

Yet, there remain to be Nash equilibria in mixed strategies which are much more efficient than the classical outcome of the equilibrium in dominant strategies \( P_A(D, D) = P_B(D, D) = 1 \). In a mixed strategy of Alice, say, she selects a particular quantum strategy \( s_A \) (which is also called pure strategy) from the set of strategies \( S_A \) with a certain classical probability. She selects the strategy \( s_A^{(i)} \) (which is also called pure strategy) out of a set of quantum strategies \( s_A^{(1)}, ..., s_A^{(N)} \), \( i = 1, ..., N \), with probability \( p_A^{(i)} \in [0, 1] \), where \( p_A^{(i)} \in [0, 1] \) and \( \sum_{i=1}^{N} p_A^{(i)} = 1 \).

One can also allow for mixed strategies by specifying a set \( s_A^{(1)}, ..., s_A^{(N)} \) of non-randomized quantum strategies and probabilities \( p_A^{(1)}, ..., p_A^{(N)} \) with \( p_A^{(i)} \in [0, 1], i = 1, ..., N, \) and \( \sum_{i=1}^{N} p_A^{(i)} = 1 \). Alice then selects the strategy (which is also called pure strategy) \( s_A^{(i)} \) with the classical probability \( p_A^{(i)} \).

That is, mixed strategies of Alice and Bob are associated with maps of the form
\[ \rho \rightarrow \sigma = \sum_{i,j} p_A^{(i)} p_B^{(j)} (U_A^{(i)} \otimes U_B^{(j)}) \rho (U_A^{(i)} \otimes U_B^{(j)})^\dagger, \quad (6.12) \]
where \( p_A^{(i)}, p_B^{(j)} \in [0, 1], i, j = 1, 2, ..., N, \) with \( \sum_i p_A^{(i)} = \sum_j p_B^{(j)} = 1 \). \( U_A^{(i)} \) and \( U_B^{(j)} \) are local unitary operators acting in \( \mathcal{H}_A \) only. corresponding to pure strategies \( s_A^{(i)} \) and \( s_B^{(j)} \).
The map given by (6.12) acts in $\mathcal{H}_A$ and $\mathcal{H}_A$ as a doubly stochastic map, that is, as a completely positive unital map [Alberti and Uhlmann (1982)]. As a result, by Uhlmann’s theorem [Marshall and Olkin (1979)] the final reduced states $\text{tr}_B[\sigma]$ and $\text{tr}_A[\sigma]$ must be more mixed than the reduced initial states $\text{tr}_B[\rho]$ and $\text{tr}_A[\rho]$ in the sense of majorisation theory [Marshall and Olkin (1979)]. As the initial state $\rho$ is a maximally entangled state, all accessible states after application of a mixed strategy of Alice and Bob are locally identical to the maximally mixed state $\frac{1}{\dim(\mathcal{H}_A)} = \frac{1}{\dim(\mathcal{H}_B)}$, which is a multiple of 1.

A particular Nash equilibrium in mixed strategies can easily be identified, more than one Nash equilibrium can be found in this quantum game. The following construction, e.g., yields an equilibrium in mixed quantum strategies: Allow Alice to choose from two strategies $s^{(1)}_A$ and $s^{(2)}_A$ with probabilities $p^{(1)}_A = 1/2$ and $p^{(2)}_A = 1/2$, while Bob may take $s^{(1)}_B$ or $s^{(2)}_B$, with

$$s^{(1)}_A \sim \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad s^{(2)}_A \sim \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$

$$s^{(1)}_B \sim \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad s^{(2)}_B \sim \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}.$$  

(6.13) \hspace{1cm} (6.14)

His probabilities are also given by $p^{(1)}_B = 1/2$ and $p^{(2)}_B = 1/2$. The quantum strategies of (6.13) and (6.14) are mutually optimal answers and have the property that

$$P_A(s^{(i)}_A, s^{(i)}_B) = 0, \quad P_B(s^{(i)}_A, s^{(i)}_B) = 5, \quad \text{for } i = 1, 2.$$  

(6.15) \hspace{1cm} (6.16)

for $i = 1, 2$.

Such quantum strategies always exist (as implied by (6.11)). Take, e.g.,

The continuous set can be achieved by choosing an arbitrary local operation $s^{(1)}_A$ and by designing $s^{(2)}_A$, $s^{(1)}_B$, and $s^{(2)}_B$ according to (6.15) and (6.16).

Due to the particular constraints of (6.15) and (6.16), there exists no other mixed strategy for Bob yielding a better pay–off than the above mixed strategy, given that Alice sticks to the equilibrium strategy. This can be seen as follows. Let Alice use this particular mixed quantum strategy as above and let Bob use any mixed quantum strategy

$$s^{(1)}_B, ..., s^{(N)}_B.$$  

(6.17)
together with \( p_{(1)}^A, ..., p_{(N)}^A \). The final state \( \sigma \) after application of the strategies is given by the convex combination

\[
\sigma = \sum_{i=1}^{2} \sum_{j} p_{(i)}^A p_{(j)}^B (s_{(i)}^A \otimes s_{(j)}^B)(\rho),
\]

(6.18)

This convex combination can only lead to a smaller expected pay-off for Bob than the optimal pure strategy \( s_{(k)}^B \) in (6.17), \( k \in \{1, ..., N\} \). Such optimal pure strategies are given by \( s_{(1)}^B \) and \( s_{(2)}^B \) as in (6.14) leading to an expected pay-off for Bob of \( P_B(s_A, s_B) = 2.5 \); there are no pure strategies which achieve a larger expected pay-off. While both pure strategies \( s_{(1)}^B \) and \( s_{(2)}^B \) do not correspond to an equilibrium, the mixed strategy where \( s_{(1)}^B \) and \( s_{(2)}^B \) are chosen with \( p_{(1)}^B = 1/2 \) and \( p_{(2)}^B = 1/2 \) actually does. Nash equilibria consist of pairs of mutually optimal answers, and only for this choice of Bob the original mixed quantum strategy of Alice is her optimal choice, as the same argument applies also to her, the game being symmetric.

This Nash equilibrium is however not the only one. There exist also other four-tuples of matrices than the ones presented in (6.13) and (6.14) that satisfy (6.15) and (6.16). Such matrices can be made out by appropriately rotating the matrices of (6.13) and (6.14). In the light of the fact that there is more than one equilibrium it is not obvious which Nash equilibrium the players will realise. It is at first not even evident whether a Nash equilibrium will be played at all. But the game theoretical concept of the focal point effect [Schelling (2060); Myerson (1991)] helps to resolve this issue.

To explore the general structure of any Nash equilibrium in mixed strategies we continue as follows: let

\[
U_{(1)}^A, ..., U_{(N)}^A
\]

(6.19)

together with \( p_{(1)}^A, ..., p_{(N)}^A \) specify the mixed strategy pertinent to a Nash equilibrium of Alice. Then there exists a mixed strategy \( U_{(1)}^B, ..., U_{(N)}^B, p_{(1)}^B, ..., p_{(N)}^B \) of Bob which rewards Bob with the best achievable pay-off, given that Alice plays this mixed strategy. Yet, the pair of mixed strategies associated with

\[
QU_{(1)}^A Q^\dagger, ..., QU_{(N)}^A Q^\dagger, QU_{(1)}^B Q^\dagger, ..., QU_{(N)}^B Q^\dagger
\]

(6.20)

with \( p_{(1)}^A, ..., p_{(N)}^A \), \( p_{(1)}^B, ..., p_{(N)}^B \) is another Nash equilibrium. This equilibrium leads to the same expected pay-off for both players, and is fully symmetric to the previous one. Doubly applying \( Q \) as
$QQU_A^{(1)}Q_A^†, ..., QQU_A^{(N)}Q_A^†$ results again into a situation with equivalent strategies as the original ones. For a given Nash equilibrium as above the one specified by (6.20) will be called dual equilibrium.

However, there is a single Nash equilibrium $(R, R)$ which is the only one which gives an expected pay–off of $P_A(R, R) = P_B(R, R) = 2.25$ and which is identical to its dual equilibrium: it is the simple map

$$\rho \mapsto \sigma = \frac{1}{\dim(H)}.$$  \hfill{(6.21)}

Indeed, there exist probabilities $p_A^{(1)}, ..., p_A^{(N)}$ and unitary operators $U_A^{(1)}, ..., U_A^{(N)}$ such that [Marshall and Olkin (1979)]

$$\sum_i p_A^{(i)}(U_A^{(i)} \otimes 1)\rho(U_A^{(i)} \otimes 1)^† = \frac{1}{\dim(H)}.$$  \hfill{(6.21)}

If Alice has already selected $s_A = R$, the application of $s_B = R$ will not change the state of the quantum system any more.

Assume that (6.19) and (6.20) are associated with equivalent quantum strategies. This means that they have to produce the same expected pay–off for all quantum strategies $s_B$ of Bob. If Alice and Bob apply $s_A \otimes s_B$ they get an expected pay–off according to (6.8) and (6.9); if Alice after implementation of $s_A$ manipulates the quantum system by applying the local unitary operator $Q \otimes 1$, they obtain

$$P'_A(s_A, s_B) = A_D \text{tr}[\pi_{CC}\sigma] + A_D \text{tr}[\pi_{CD}\sigma] + A_D \text{tr}[\pi_{DD}\sigma],$$  

$$P'_B(s_A, s_B) = B_D \text{tr}[\pi_{CC}\sigma] + B_D \text{tr}[\pi_{CD}\sigma] + B_D \text{tr}[\pi_{DD}\sigma].$$

The only $s_A$ with the property that $P'_A(s_A, s_B) = P_A(s_A, s_B)$ and $P'_B(s_A, s_B) = P_B(s_A, s_B)$ for all $s_B$ is the map given by (6.21).

In principle, any Nash equilibrium may become a self–fulfilling prophecy if the particular Nash equilibrium is expected by both players. It has been pointed out that in a game with more than one equilibrium, anything that attracts the players’ attention towards one of the equilibria may make them expect and therefore realise it [Schelling (2060)]. The corresponding focal equilibrium [Myerson (1991)] is the one which is conspicuously distinguished from the other Nash equilibria. In this particular situation there is indeed one Nash equilibrium different from all the others: it is the one which is equivalent to its dual equilibrium, the map which simply maps
the initial state on the maximally mixed state. For all other expected payoffs both players are ambivalent between (at least) two symmetric equilibria. The expected pay-off the players will receive in this focal equilibrium, $P_A(R,R) = P_B(R,R) = 2.25$, is not fully Pareto optimal, but it is again much more efficient than the classically achievable outcome of 1.\footnote{In the classical game both players could also play $C$ and $D$ with probabilities $p = 1/2$ yielding the same expected pay-off of 2.25 for both players, but this pair of mixed strategies would be no equilibrium solution, as any player could benefit from simply choosing the dominant strategy $D$.}

**Completely positive trace-preserving maps corresponding to local operations**

In this scenario both Alice and Bob may perform any operation that is allowed by quantum mechanics. That is, the set of strategies $S^{(CP)}$ is made up of $(s_A, s_B)$, where both $s_A$ and $s_B$ correspond to a completely positive trace-preserving map

$$
(s_A \otimes s_B)(\rho) = \sum_i \sum_j (A_i \otimes B_j)\rho(A_i \otimes B_j)^\dagger,
$$
corresponding to a local operation, associated with Kraus operators $A_i$ and $B_j$ with $i, j = 1, 2, \ldots$. The trace-preserving property requires that

$$
\sum_i A_i^\dagger A_i = 1 \quad \text{and} \quad \sum_i B_i^\dagger B_i = 1.
$$

This case has already been mentioned in [Eisert et al. (1999)]. The quantum strategies $s_A$ and $s_B$ do no longer inevitably act as unital maps in the respective Hilbert spaces as before. In other words, the reduced states of Alice and Bob after application of the quantum strategy are not necessarily identical to the maximally mixed state $1/\dim(H_A)$.

As already pointed out in [Eisert et al. (1999)], the pair of strategies $(Q, Q)$ of the two-parameter set of strategies $S^{(TP)}$ is again no equilibrium solution. It is straightforward to prove that the Nash equilibria of the type of (6.13) and (6.14) of mixed strategies with general local unitary operations are, however, still present, and each of these equilibria yields an expected pay-off of 2.5. In the language of this subsection these equilibria are equilibria in pure strategies, as mixing is already included in this most general case.

In addition, as strategies do no longer have to be locally unital maps, it is not surprising that new Nash equilibria emerge: Alice and Bob may,
e.g., perform a measurement associated with Kraus operators

\[ A_1 = |0\rangle\langle 0|, \quad A_2 = |1\rangle\langle 1|, \quad B_1 = D|0\rangle\langle 0|, \quad B_2 = D|1\rangle\langle 1|. \]

This operation yields a final state

\[ \sigma = (s_A \otimes s_B)(\rho) = (|01\rangle\langle 01| + |10\rangle\langle 10|)/2. \]

Clearly neither Alice nor Bob can gain from unilaterally deviating from their strategy. Another Nash equilibrium is associated with Kraus operators

\[ A_1 = |1\rangle\langle 1|, \quad A_2 = D|0\rangle\langle 0|, \quad B_1 = |1\rangle\langle 1|, \quad B_2 = D|0\rangle\langle 0|. \]

One can nevertheless argue as in the previous case. Again, all Nash equilibria occur at least in pairs. First, there are again the dual equilibria from \( S^{(GU)} \). Second, there are Nash equilibria \((s_A, s_B), s_A \neq s_B\), with the property that \((s_B, s_A)\) is also a Nash equilibrium yielding the same expected pay-off. The only Nash equilibrium invariant under application of \( Q \) and exchange of the strategies of the players is again \((R, R)\) defined in the previous subsection, which yields a pay-off \( P_A(R, R) = P_B(R, R) = 2.25 \). \((R, R)\) is still a Nash equilibrium, because – given that one player picks \( R \) – the application of any strategy taken from \( S^{(CP)} \) of the other player results in a final state which yields a pay-off smaller or equal to 2.25. This is the solution of the game is the most general case. While both players could in principle do better (as the solution lacks Pareto optimality), the efficiency of this focal equilibrium is much higher than the equilibrium in dominant strategies of the classical game. Hence, also in this most general case both players gain from using quantum strategies.

This study shows that the efficiency of the equilibrium the players can reach in this game depends on the actions the players may take. One feature, however is present in each of the considered sets: both players can increase their expected pay-offs drastically by resorting to quantum strategies.

### 6.2.3.2 Chicken Game

In the previous classical game, the Prisoners’ Dilemma, an unambiguous solution can be specified consisting of a unique Nash equilibrium. However, this solution was not efficient, thus giving rise to the dilemma. In the so-
called Chicken game [Myerson (1991); Poundstone (1992)],

\[
A_{CC} = B_{CC} = 6, \quad A_{CD} = B_{DC} = 8, \\
A_{DC} = B_{CD} = 2, \quad A_{DD} = B_{DD} = 0.
\]

This game has two Nash equilibria \((C, D)\) and \((D, C)\): it is not clear how to anticipate what the players’ decision would be. In addition to the two Nash equilibria in pure strategies there is an equilibrium in mixed strategies, yielding an expected pay–off 4 [Myerson (1991)].

In order to investigate the new features of the game if superpositions of classical strategies are allowed for, three set of strategies are briefly discussed:

1. One–Parameter Set of Strategies

Again, we consider the set of strategies \(S^{(CL)}\) of one-dimensional rotations. The strategies \(s_A\) and \(s_B\) are associated with local unitary operators

\[
U(\theta) = \begin{pmatrix}
\cos(\theta/2) & \sin(\theta/2) \\
-\sin(\theta/2) & \cos(\theta/2)
\end{pmatrix},
\]

with \(\theta \in [0, \pi]\),

\[
C \sim U(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad D \sim U(\pi) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.
\]

Then as before, the quantum game yields the same expected pay–off as the classical game in randomized strategies. This means that still two Nash equilibria in pure strategies are present.

2. Two–Parameter Set of Strategies

The players can actually take advantage of an additional degree of freedom which is not accessible in the classical game. If they may apply unitary operations from \(S^{(TP)}\) of the type

\[
U(\theta, \phi) = \begin{pmatrix}
e^{i\phi}\cos(\theta/2) & \sin(\theta/2) \\
-\sin(\theta/2) & e^{-i\phi}\cos(\theta/2)
\end{pmatrix},
\]

with \(\theta \in [0, \pi]\) and \(\phi \in [0, \pi/2]\) the situation is quite different than with \(S^{(CL)}\). \((C, D)\) and \((C, D)\) with \(C \sim U(0, 0)\) and \(D \sim U(\pi, 0)\) are no longer equilibrium solutions. E.g., given that \(s_A = D\) the pair of strategies \((D, Q)\) with \(Q \sim U(0, \pi/2)\) yields a better expected pay–off for Bob than \((D, C)\), that is to say \(P_B(D, Q) = 8, P_B(D, C) = 2\). In fact \((Q, Q)\) is now the unique
Nash equilibrium with $P_A(Q, Q) = P_B(Q, Q) = 6$, which follows from an investigation of the actual expected pay-offs of Alice and Bob analogous to (6.10). This solution is not only the unique acceptable solution of the game, but it is also an equilibrium that is Pareto optimal. This contrasts very much with the situation in the classical game, where the two equilibria were not that efficient.

3. Completely Positive Trace–Preserving Maps

As in the considerations concerning the Prisoner’s Dilemma game, more than one Nash equilibrium is present, if both players can take quantum strategies from the set $S^{(CP)}$, and all Nash equilibria emerge at least in pairs as above. The focal equilibrium is given by $(R, R)$, resulting in a pay-off of $P_A(R, R) = P_B(R, R) = 4$, which is the same as the mixed strategy of the classical game.

6.2.4 Quantum Cryptography and Gambling

Recall that quantum cryptography is a field which combines quantum theory with information theory. The goal of this field is to use the laws of physics to provide secure information exchange, in contrast to classical methods based on (unproven) complexity assumptions. In particular, quantum key distribution protocols [Ekert (1991)] became especially important due to technological advances which allow their implementation in the laboratory. However, the last important theoretical result in the field was of a negative character: Mayers [Mayers (1997)] and Lo and Chau [Lo and Chau (1997)] showed that quantum bit commitment is not secure. Their work also raised serious doubts on the possibility of obtaining any secure two–party protocol, such as oblivious transfer and coin tossing [Lo (1997)]. In this subsection, following [Goldenberg et al. (1999)] we present a secure two–party quantum cryptographic task – quantum gambling, which has no classical counterpart.

Coin tossing is defined as a method of generating a random bit over a communication channel between two distant parties. The parties, traditionally named Alice and Bob, do not trust each other, or a third party. They create the random bit by exchanging quantum and classical information. At the end of the protocol the generated bit is known to both of them. If a party cheats, i.e. changes the occurrence probability of an outcome, the other party should be able to detect the cheating. We would consider a coin tossing protocol to be secure if it defines a parameter such that when
it goes to infinity the probability to detect any finite change of probabilities goes to 1. Using a secure protocol the parties can make certain decisions depending on the value of the random bit, without being afraid that the opponent may have some advantage. For instance, Alice and Bob can play a game in which Alice wins if the outcome is '0' and Bob wins if it is '1'. Note that if bit commitment were secure, it could be used to implement coin tossing trivially: Alice commits a bit \( a \) to Bob; Bob tells Alice the value of a bit \( b \); the random bit is the parity bit of \( a \) and \( b \).

It is not known today if a secure quantum coin tossing protocol can be found.\(^{17}\) It is only known that ideal coin tossing, i.e., in which no party can change the expected distribution of the outcomes, is impossible [Lo and Chau (1997)]. Based on our efforts in this direction, we are skeptical about the possibility to have secure (non–ideal) coin tossing. Nevertheless, we were able to construct a protocol which gives a solution to a closely related task. “Quantum gambling” is very close to playing in a casino located in a remote site, such as gambling over the Internet. As in a real casino, for instance when playing Roulette, the player’s possible choices give him some probability to win twice the amount of his bet, or a smaller probability to win a bigger sum. However, in our protocol the player has only a partial control over these choices. In spite of its limitations our protocol provides a quantum solution to a useful task, which cannot be performed securely today in the classical framework. Assuming ideal apparatus and communication channels, the protocol is unconditionally secure, depending solely on the laws of physics.

Let us start by defining exactly the gambling task considered here. The casino (Alice) and the player (Bob) are physically separated, communicating via quantum and classical channels. The bet of Bob in a single game is taken for simplicity to be 1 coin. At the end of a game the player wins 1 or \( R \) coins, or loses 1 coin (his bet), depending on the result of the game. We have found a protocol which implement this game while respecting two requirements: First, the player can ensure that, irrespective of what the casino does, his expected gain is not less than \( \delta \) coins, where \( \delta \) is a negative function of \( R \) which goes to zero when \( R \) goes to infinity. The exact form of \( \delta(R) \) will be specified below. Second, the casino can ensure that, irrespective of what the player does, its expected gain is not less than 0 coins.

\(^{17}\)If we limit ourselves to spatially extended secure sites located one near the other, then secure coin tossing can be realized classically, by simultaneous exchange of information at the opposite sides of the sites. The security of this method relies on relativistic causality.
In order to define the protocol rigorously, we will first present the rules of the game, then the strategies of the players which ensure the outcomes quoted above and finally we will prove the security of the method.

**The Rules of the Game:** Alice has two boxes, $A$ and $B$, which can store a particle [Goldenberg *et al.* (1999)]. The quantum states of the particle in the boxes are denoted by $|a⟩$ and $|b⟩$, respectively. Alice prepares the particle in some state and sends box $B$ to Bob.

Bob wins in one of the two cases:

1. If he finds the particle in box $B$, then Alice pays him 1 coin (after checking that box $A$ is empty).
2. If he asks Alice to send him box $A$ for verification and he finds that she initially prepared a state different from

$$|ψ_0⟩ = \frac{1}{\sqrt{2}}(|a⟩ + |b⟩), \quad (6.22)$$

then Alice pays him $R$ coins.

In any other case Alice wins, and Bob pays her 1 coin.

The players’ strategies which ensure (independently) an expectation value of Alice’s gain $G_A \geq 0$ (irrespective of Bob’s actions) and an expectation value of Bob’s gain $G_B \geq δ$ (irrespective of Alice’s actions) are as follows:

**Alice’s Strategy:** Alice prepares the equally distributed state $|ψ_0⟩$ (given in (6.22)).

**Bob’s Strategy:** After receiving box $B$, Bob splits the particle in two parts; specifically, he performs the following unitary operation:

$$|b⟩ \rightarrow \sqrt{1 - η} |b⟩ + \sqrt{η} |b′⟩, \quad (6.23)$$

where $⟨b′|b⟩ = 0$. The particular splitting parameter $η$ he uses is $η = ˜η(R)$ (to be specified below). After the splitting Bob measures the projection operator on the state $|b⟩$, and then

I. If the measurement yields a positive result, i.e. he finds the particle, he announces Alice that he won.

II. If the measurement yields a negative result, he asks Alice for box $A$ and verifies the preparation.

This completes the formal definition of our protocol.
In order to prove the security of the scheme, we will analyze the average gain of each party as a result of her/his specific strategy. It is straightforward to see that Alice’s strategy ensures $G_A \geq 0$. If Alice prepares the state $|\psi_0\rangle$, Bob has no meaningful way of increasing his odds beyond 50%: if he decides to open box $B$ he has a probability of 0.5 to win 1 coin and a probability of 0.5 to lose 1 coin. He cannot cheat by claiming that he found the particle when he did not, since Alice learns the result by opening box $A$. If, instead, he decides to verify the preparation he will find the expected state, so he will lose 1 coin. Therefore $G_B \leq 0$, and since this is a zero–sum game, Alice’s gain is $G_A \geq 0$, whatever Bob does.

Now we will prove that Bob, using the splitting parameter $\eta = \tilde{\eta}$, can ensure $G_B \geq \delta$. The values of $\tilde{\eta}$ and $\delta$ are determined by the calculation of Bob’s expected gain, $G_B$. Bob tries to maximize $G_B$ under the assumption that Alice uses the worse strategy for him, namely the one which minimizes $G_B$ for Bob’s particular strategy. Therefore, we will first minimize the function $G_B$ for any $\eta$, and then we will find the maximum of the obtained function, with that computing $\delta$. We will also compute the value of $\eta$ at the peak, $\tilde{\eta}$, which will be the chosen splitting parameter of Bob.

Let us first write down the expression for $G_B$. Bob gets 1 coin if he detects the state $|b\rangle$; denote the probability for this event to occur by $P_b$. He gets $R$ coins if he detects a different preparation than $|\psi_0\rangle$ (after failing to find the state $|b\rangle$, an event with a related probability of $1 - P_b$); denote the probability to detect a different preparation by $P_D$. He loses 1 coin if he does not detect a different preparation than $|\psi_0\rangle$ (after failing to find $|b\rangle$); the probability for this event is $(1 - P_D)$. Thus, the expectation value of Bob’s gain is

$$G_B = P_b + (1 - P_b) [P_D R - (1 - P_D)].$$

(6.24)

For the calculations of $P_b$ and $P_D$ we will consider the most general state Alice can prepare. In this case the particle may be located not only in boxes $A$ and $B$, but also in other boxes $C_i$. The states $|a\rangle$, $|b\rangle$ and $|c_i\rangle$ are mutually orthogonal. She can also correlate the particle to an ancilla $|\Phi\rangle$, such that the most general preparation is

$$|\Psi_0\rangle = \alpha|a\rangle|\Phi_a\rangle + \beta|b\rangle|\Phi_b\rangle + \sum_i \gamma_i|c_i\rangle|\Phi_{c_i}\rangle,$$

(6.25)

where $|\Phi_a\rangle$, $|\Phi_b\rangle$, $|\Phi_{c_i}\rangle$ are the states of the ancilla and $|\alpha|^2 + |\beta|^2 + \sum_i |\gamma_i|^2 = 1$. After Bob splits $|b\rangle$, as described by (6.23), the state changes
to

$$|\Psi_1\rangle = \alpha|a\rangle|\Phi_a\rangle + \beta \left( \sqrt{1-\eta} |b\rangle + \sqrt{\eta} |b'\rangle \right) |\Phi_b\rangle + \sum_i \gamma_i |c_i\rangle |\Phi_{c_i}\rangle. \quad (6.26)$$

The probability to find the state $|b\rangle$ (in step I. of Bob’s strategy) is

$$P_b = \|\langle b|\Psi_1\rangle\|^2 = |\beta|^2 (1-\eta). \quad (6.27)$$

If Bob does not find $|b\rangle$, then the state reduces to

$$|\Psi_2\rangle = N \left( \alpha|a\rangle|\Phi_a\rangle + \beta \sqrt{\eta} |b'\rangle |\Phi_b\rangle + \sum_i \gamma_i |c_i\rangle |\Phi_{c_i}\rangle \right), \quad (6.28)$$

where $N$ is the normalization factor given by $N = (1 - (1-\eta) |\beta|^2)^{-1/2}$.

On the other hand, if Alice prepares the state $|\psi_0\rangle$ instead of $|\Psi_0\rangle$, then at this stage the particle is in the state

$$|\psi_2\rangle = \sqrt{\frac{1}{1+\eta}} |a\rangle + \sqrt{\frac{\eta}{1+\eta}} |b'\rangle. \quad (6.29)$$

Thus, the best verification measurement of Bob is to make a projection measurement on this state. If the outcome is negative, Bob knows with certainty that Alice did not prepare the state $|\psi_0\rangle$. The probability of detecting such a different preparation is given by

$$P_D = 1 - \|\langle \psi_2|\Psi_2\rangle\|^2 = 1 - N^2 \left( \frac{\alpha}{\sqrt{1+\eta}} |\Phi_a\rangle + \frac{\beta \eta}{\sqrt{1+\eta}} |\Phi_b\rangle \right)^2. \quad (6.30)$$

Since Alice wants to minimize $G_B$, she tries to minimize both $P_b$ and $P_D$. From (6.30) we see that in order to minimize $P_D$, the states of the ancilla $|\Phi_a\rangle$ and $|\Phi_b\rangle$ have to be identical (up to some arbitrary phase), i.e., $|\langle \Phi_a|\Phi_b\rangle| = 1$. That is, Alice gets no advantage using an ancilla, so it can be eliminated. Then, in order to maximize $N|\alpha + \beta \eta|$, Alice should set all $\gamma_i$ to zero, as it is clear from the normalization constraint

$$|\alpha|^2 + |\beta|^2 = 1 - \sum_i |\gamma_i|^2.$$

This operation has no conflict with the minimization of $P_b$, since (6.27) contains only $|\beta|$. Also, the maximization is possible if the coefficients $\alpha$ and $\beta$, if seen as vectors in the complex space, point in the same direction. Therefore, Alice gains nothing by taking $\alpha$ and $\beta$ to be complex numbers; it
Quantum Leap

is sufficient to use real positive coefficients. Taking all these considerations into account, the state prepared by Alice can be simplified to

$$|\psi'_0\rangle = \sqrt{\frac{1}{2} + \epsilon} |a\rangle + \sqrt{\frac{1}{2} - \epsilon} |b\rangle.$$  \hfill (6.31)

Now, the state after Bob splits $|b\rangle$ reads

$$|\psi'_1\rangle = \sqrt{\frac{1}{2} + \epsilon} |a\rangle + \sqrt{\frac{1}{2} - \epsilon} \left( \sqrt{1 - \eta} |b\rangle + \sqrt{\eta} |b'\rangle \right),$$  \hfill (6.32)

and so the probability to find $|b\rangle$ becomes

$$P_b = ||\langle b|\psi'_1\rangle||^2 = \left( \frac{1}{2} - \epsilon \right) (1 - \eta).$$  \hfill (6.33)

When Bob does not find the state $|b\rangle$, $|\psi'_1\rangle$ reduces to

$$|\psi'_2\rangle = \frac{\sqrt{1 + 2\epsilon} |a\rangle + \sqrt{\eta (1 - 2\epsilon)} |b'\rangle}{\sqrt{1 + 2\epsilon + \eta (1 - 2\epsilon)}},$$  \hfill (6.34)

which in turn leads to

$$P_D = 1 - ||\langle \psi_2|\psi'_2\rangle||^2 = \frac{2\eta \left( 1 - \sqrt{1 - 4\epsilon^2} \right)}{(1 + \eta)^2 + 2\epsilon (1 - \eta^2)}.$$  \hfill (6.35)

Substituting (6.33) and (6.35) in (6.24), we find $G_B$ in terms of the splitting parameter $\eta$, the preparation parameter $\epsilon$ and $R$:

$$G_B = -\frac{1}{1 + \eta} \left[ 2\epsilon (1 - \eta^2) + \eta (\eta + \sqrt{1 - 4\epsilon^2}) 
- \eta (1 - \sqrt{1 - 4\epsilon^2} \cdot R) \right].$$  \hfill (6.36)

In order to calculate the minimal gain of Bob, $\delta$, irrespective of the particular strategy of Alice, we will first minimize $G_B$ for $\epsilon$ and then maximize the result for $\eta$:

$$\delta(R) = \text{Max}_{\eta} [\text{Min}_{\epsilon} G_B(R, \eta, \epsilon)].$$  \hfill (6.37)
The calculations yield
\[
\delta = -\frac{1}{1 + \sqrt{R + 2 - \sqrt{(R + 2)^2 - 1}}}
\times \left\{ 2 + \left[ R - \sqrt{(R + 2)^2 - 1} \right] \right\} \times \left[ 1 - \sqrt{R + 2 - \sqrt{(R + 2)^2 - 1}} \right],
\]
(6.38)

obtained for Bob's splitting parameter
\[
\tilde{\eta} = \sqrt{R + 2 - \sqrt{(R + 2)^2 - 1}}.
\]
(6.39)

In the range of \( R \gg 1 \), these results can be simplified to
\[
\delta \approx -\sqrt{\frac{2}{R}},
\]
(6.40)
\[
\tilde{\eta} \approx \sqrt{\frac{1}{2R}}.
\]
(6.41)

We have shown that if Bob follows the proposed strategy with \( \eta = \tilde{\eta} \), then his average gain is not less than \( \delta \); this bound converges to 0, i.e. to the limit of a fair game, for \( R \to \infty \). This is true for any possible strategy of Alice, therefore, the security of the protocol is established [Goldenberg et al. (1999)].

To compare our scheme to a real gambling situation, let us consider the well-known Roulette game. A bet of 1 coin on the red/black numbers, i.e. half of the 36 numbers on the table, rewards the gambler with 1 coin once in 18/38 turns (on average, for a spinning wheel with 38 slots); this gives an expected gain of about \(-0.053\) coins. To assure the same gain in our scheme, \( R = 700 \) is required. Note that extremely large values of \( R \) are practically meaningless, one reason being the limited total amount of money in use. Nevertheless, the bound on \( \delta \) is not too restrictive when looking at the first prizes of some lottery games: a typical value of \( R = 10^6 \) gives a reasonably small \( \delta \) of about \(-0.0014\).

It is also interesting to consider the case of \( R = 1 \). This case corresponds to coin tossing, since it has only two outcomes: Bob’s gain is either \(-1\) coin (stands for bit ‘0’) or 1 coin (stands for bit ‘1’). The minimal average gain of Bob is about \(-0.657\), which translates to an occurrence probability of bit ‘1’ of at least 0.172 (instead of 0.5 ideally), whatever Alice does. This is certainly not a good coin tossing scheme, however, no classical or quantum
method is known to assure (unconditionally) any bound for the occurrence probability of both outcomes.

Our analysis so far was restricted to a single instance of the game, but the protocol may be repeated several times. After $N$ games Bob’s expected gain is $G_B \geq N \delta$ and Alice’s expected gain is $G_A \geq 0$. Of course, Alice may choose now a complex strategy using ancillas and correlations between particles/ancillas from different runs. In this way she may change the probability distribution of her winnings, but she cannot reduce the minimal expected gain of Bob. Indeed, our proof considers the most general actions of Alice, so the average gain of Bob in each game is not less than $\delta$, and consequently, it is not less then $N \delta$ after $N$ games. A similar argument is valid for Bob’s actions, so the average gain of Alice remains non-negative even after $N$ games. In gambling games, in addition to the average gain, it is important to analyze the standard deviation of the gain, $\Delta G$. Bob will normally accept to play a game with a negative gain only if $\Delta G_B \gg |G_B|$ (unless he has some specific target in mind). In a single application of our protocol, $\Delta G_B \geq 1$, so the condition is attained for big enough values of $R$ (see (6.40)). However, increasing the number of games makes the gambling less attractive to Bob: if Alice follows the proposed strategy, $|G_B|$ grows as $N$ while $\Delta G_B$ grows only as $\sqrt{N}$. Therefore, Bob should accept to play $N$ times only if $N \ll 1/\delta^2 \sim R$.

Another important point to consider is the possible “cheating” of the parties. Alice has no meaningful way to cheat, since she is allowed to prepare any quantum state and she sends no classical information to Bob. Any operation other than preparing $|\psi_0\rangle$, as adding ancillas or putting more/less than one particle in the boxes, just decreases her minimal gain. Bob, however, may try to cheat. He may claim that he detected a different preparation than $|\psi_0\rangle$, even when his verification measurement does not show that. If Alice prepares the initial state $|\psi_0^{\prime}\rangle$ (with $\epsilon > 0$), she is vulnerable to this cheating attempt: she has no way to know if Bob is lying or not. For this reason Alice’s proposed strategy is to prepare $|\psi_0\rangle$ every time, such that any cheating of Bob could be invariably detected. When both parties follow the proposed strategies, i.e. $\epsilon = 0$ and $\eta = \tilde{\eta}$, the game is more fair for Bob than assumed in the proof [Goldenberg et al. (1999)]:

$$G_{B_{\text{prot}}} = -G_{A_{\text{prot}}} = -\sqrt{R + 2 - \sqrt{(R + 2)^2 - 1}}. \quad (6.42)$$

For $R \gg 1$ we get $G_{B_{\text{prot}}} \approx -1/\sqrt{2R}$, which is approximately half of the value of $\delta$ calculated in (6.40).
The discussion up to this point assumed an ideal experimental setup. In practice errors are unavoidable, of course, and our protocol is very sensitive to the errors caused by the devices used in its implementation (communication channels, detectors, etc). In the presence of errors, if the parties disagree about the result of a particular run it should be canceled. If such conflicts occur more than expected based on the experimental error rate, it means that (at least) one party is cheating, and the game should be stopped. The most sensitive part to errors is the verification measurement of Bob, i.e. the detection of the possible deviation of the initial state from $|\psi_0\rangle$. In the ideal case, using $\tilde{\eta}$ and the corresponding $\epsilon$ (the worst for honest Bob), the detection probability is very small: $P_D \approx \sqrt{2/R^3}$, for $R \gg 1$. Clearly, for a successful realization of the protocol, the error rate has to be lower than this number. Thus, in practice, the experimental error rate will constrain the maximal possible value of $R$.\footnote{In the presence of errors an alternative strategy of Bob, without splitting state $|b\rangle$, is more efficient: Bob requests box $A$ for verification randomly (at a rate optimized for the value of $R$) and measures the projection on $|\psi_0\rangle$ (not on $|\psi_2\rangle$). Although for a given $R$ this strategy offers him a lower gain ($\delta \approx -2/\sqrt{R}$, for $R \gg 1$), it allows using a higher $R$ for a given fidelity of the setup, since now $P_D \approx 1/R$.}

\subsection*{6.3 Formal Quantum Games}

Recall that Meyer [Meyer and Schmidt (2000)] and Eisert et al. [Eisert et al. (1999)] brought the game theory into the physics community and created a new field, quantum game theory. They both quantized a classical game and found interesting new properties which the original classical game does not possess. Nevertheless, their quantized games seem quite different. PQ penny flipover studied by Meyer is a quantum sequential game, in which players take turns in performing some operations on a quantum system. On the other hand, quantum Prisoners’ Dilemma studied by Eisert et al. is a quantum simultaneous game, in which there are $n$ players and a quantum system which consists of $n$ subsystems, and player $i$ performs an operation only on the $i$th subsystem.

Since the seminal works of Meyer and Eisert et al., many studies have been made to quantize classical games and find interesting phenomena [Eisert and Wilkens (2000); Marinatto and Weber (2000); Benjamin and Hayden (2001); Du (2002); Flitney and Abbott (2002); Flitney and Abbott (2004)]. Most of the quantum games ever studied are classified into either quantum simultaneous games or quantum sequential...
games, although not much has been done on the latter.

Now that we see that game theory is combined with quantum theory and there are two types of quantum games, several questions naturally arise: (a) Are quantum games truly different from classical games? (b) If so, in what sense are they different? (c) What is the relationship between quantum simultaneous games and quantum sequential games? To answer these questions, it is necessary to examine the whole structure of game theory including classical games and quantum games, not a particular phenomenon of a particular game.

A work by Lee and Johnson [Lee and Johnson (2003)] is a study along this line. They developed a formalism of games including classical games and quantum games. With the formalism they addressed the questions (a) and (b), concluding that “playing games quantum mechanically can be more efficient” and that “finite classical games consist of a strict subset of finite quantum games”. However, they did not give a precise definition of the phrase “consist of a strict subset”.

In this section, following [Kobayashi (2007)], we present various types of quantum games and try to answer the above questions.

### 6.3.1 Formal Framework

In order to discuss relationships between different types of games, we need a common framework in which various types of games are described. As the first step in our analysis, we will construct such a framework for our theory.

For the construction, a good place to start is to consider what is game theory. Game theory is the mathematical study of game situations which is characterized by the following three features: (i) There are two or more decision-makers, or players; (ii) Each player develops his/her strategy for pursuing his/her objectives. On the basis of the strategy, he/she chooses his/her action from possible alternatives; and (iii) As a result of all players' actions, some situation is realized. Whether the situation is preferable or not for one player depends not only on his/her action, but also on the other players’ actions.

How much the realized situation is preferable for a player is quantified by a real number called a pay-off. Using this term, we can rephrase the second feature as “each player develops his/her strategy to maximize the expectation value of his/her pay-off”. The reason why the expectation value is used to evaluate strategies is that we can determine the resulting
situation only probabilistically in general, even when all players’ strategies are known.

As a mathematical representation of the three features of game situations, we define a normal form of a game.

A normal form of a game is a triplet \((N, \Omega, f)\) whose components satisfy the following conditions: \(N = \{1, 2, \ldots, n\}\) is a finite set; \(\Omega = \Omega_1 \times \cdots \times \Omega_n\), where \(\Omega_i\) is a nonempty set; and \(f\) is a function from \(\Omega\) to \(\mathbb{R}^n\).

Here, \(N\) denotes a set of players. \(\Omega_i\) is a set of player \(i\)’s strategies, which prescribes how he/she acts. The \(i\)th element of \(f(\omega_1, \ldots, \omega_n)\) is the expectation value of the pay–off for player \(i\), when player \(j\) adopts a strategy \(\omega_j\).

Next, we propose a general definition of games, which works as a framework for discussing relationship between various kinds of games. We can regard a game as consisting of some ‘entities’ (like players, cards, coins, etc.) and a set of rules under which a game situation occurs. We model the ‘entities’ in the form of a tuple \(T\). Furthermore, we represent the game situation caused by the ‘entities’ \(T\) under a rule \(R\) as a normal form of a game, and write it as \(R(T)\). Using these formulations, we define a game as follows.

We define a game as a pair \((T, R)\), where \(T\) is a tuple, and \(R\) is a rule which determines uniquely a normal form from \(T\). When \(G = (T, R)\) is a game, we refer to \(R(T)\) as the normal form of the game \(G\).

The conception of the above definition will be clearer if we describe various kinds of games in the form of the pair defined above. Thus far, we have implicitly regarded strategies and actions of individual players as elementary components of a game. In classical game theory, such modelling of games is referred to as a noncooperative game, in contrast to a cooperative game in which strategies and actions of groups of players are elementary.

### 6.3.2 Various Types of Formal Games

In this subsection, following [Kobayashi (2007)](kobayashi2007), we introduce various types of classical and quantum games. First, we confirm that strategic games, which is a well–established representation of games in classical game theory (see, e.g., [Osborne and Rubinstein (1994)](osborne1994)), can be described in our framework. Then, we define two quantum games, namely, quantum simultaneous games and quantum sequential games.
6.3.2.1 Strategic Games

We can redefine strategic games using our formal framework as follows.

A strategic game is a game \((T, R)\) which has the following form: (i) \(T = (N, S, f)\), and each component satisfies the following conditions: \(N = \{1, 2, \ldots, n\}\) is a finite set; \(S = S_1 \times \cdots \times S_n\), where \(S_i\) is a nonempty set; and \(f : S \mapsto \mathbb{R}^n\) is a function from \(S\) to \(\mathbb{R}^n\). (ii) \(R(T) = T = (N, S, f)\).

If the set \(S_i\) is finite for all \(i\), then we call the game \((T, R)\) a finite strategic game. We denote the set of all strategic games by \(SG\), and the set of all finite strategic games by \(FSG\).

Let \(G = ((N, S, f), R)\) be a finite strategic game. Then the mixed extension of \(G\) is a game \(G^* = ((N, S, f), R^*)\), where the rule \(R^*\) is described as follows: \(R^*(N, S, f) = (N, Q, F)\), where \(Q\) and \(F\) are of the following forms:

\[
Q = Q_1 \times \cdots \times Q_n,
\]

where \(Q_i\) is the set of all probability distribution over \(S_i\); \(F : Q \mapsto \mathbb{R}^n\) assigns to each \((q_1, \ldots, q_n) \in Q\) the expected value of \(f\). That is, the value of \(F\) is given by

\[
F(q_1, \ldots, q_n) = \sum_{s_1 \in S_1} \cdots \sum_{s_n \in S_n} \left\{ \prod_{i=1}^{n} q_i(s_i) \right\} f(s_1, \ldots, s_n),
\]

where \(q_i(s_i)\) is the probability attached to \(s_i\).

We denote the set of all mixed extensions of finite strategic games by \(MEFSG\).

6.3.2.2 Quantum Simultaneous Games

Quantum simultaneous games are quantum games in which a quantum system is used according to a protocol depicted in Figure 6.3. In quantum simultaneous games, there are \(n\) players who can not communicate with each other, and a referee. The referee prepares a quantum system in the
initial state $\hat{\rho}_{\text{init}}$. The quantum system is composed of $n$ subsystems, where the Hilbert space for the $i$th subsystem is $\mathcal{H}_i$. The referee provides player $i$ with the $i$th subsystem. Each player performs some quantum operation on the provided subsystem. It is determined in advance which operations are available for each player. After all players finish their operations, they return the subsystems to the referee. Then the referee performs a POVM measurement $\{\hat{M}_r\}$ on the total system. If the $r$th measurement outcome is obtained, player $i$ receives a pay–off $a^i_r$.

Many studies on quantum simultaneous games have been carried out. Early significant studies include [Eisert et al. (1999); Eisert and Wilkens (2000); Marinatto and Weber (2000); Benjamin and Hayden (2001); Du (2002)].

The protocol of the quantum simultaneous games is formulated as follows.

A quantum simultaneous game is a game $(T, R)$ which has the following form.

1. $T = (N, \mathcal{H}, \hat{\rho}_{\text{init}}, \{\hat{M}_r\}, \{a_r\})$, and each component satisfies the following conditions: (i) $N = \{1, 2, \ldots, n\}$ is a finite set; (ii) $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_n$, where $\mathcal{H}_i$ is a Hilbert space; (iii) $\hat{\rho}_{\text{init}}$ is a density operator on $\mathcal{H}$; (iv) $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$, where $\Omega_i$ is a subset of the set of all CPTP (completely positive trace preserving) maps on the set of density operators on $\mathcal{H}_i$. In other words, $\Omega_i$ is a set of quantum operations available for player $i$; (v) $\{\hat{M}_r\}$ is a POVM on $\mathcal{H}$; and (vi) $a_r = (a^1_r, a^2_r, \ldots, a^n_r) \in \mathbb{R}^n$. The index $r$ of $a_r$ runs over the same domain as that of $\hat{M}_r$.

2. $R(T) = (N, \Omega, f)$. The value of $f$ is given by

$$f(\mathcal{E}_1, \ldots, \mathcal{E}_n) = \sum_r a_r \text{Tr} \left[ \hat{M}_r (\mathcal{E}_1 \otimes \cdots \otimes \mathcal{E}_n) (\hat{\rho}_{\text{init}}) \right]$$

for all $(\mathcal{E}_1, \ldots, \mathcal{E}_n) \in \Omega$.

If $\mathcal{H}_i$ is finite dimensional for all $i$, then we refer to the game $(T, R)$ as a finite quantum simultaneous game. We denote the set of all quantum simultaneous games by $QSim$, and the set of all finite quantum simultaneous games by $FQSim$. 
6.3.2.3 Quantum Sequential Games

Quantum sequential games are another type of quantum games, in which a quantum system is used according to a protocol depicted in Figure 6.4. In quantum sequential games, there are $n$ players who can not communicate each other and a referee. The referee prepares a quantum system in the initial state $\hat{\rho}_{\text{init}}$. The players performs quantum operations on the quantum system in turn. The order of the turn may be regular like $1 \rightarrow 2 \rightarrow 3 \rightarrow \cdots$, or may be irregular like $1 \rightarrow 3 \rightarrow 2 \rightarrow 3 \rightarrow 1 \rightarrow 2 \rightarrow \cdots$, yet it is determined in advance. After all the $m$ operations are finished, the referee performs a POVM measurement $\{\hat{M}_r\}$. If the $r$-th measurement outcome is obtained, then player $i$ receives a pay–off $a_{ir}$.

Games which belong to quantum sequential games include $PQ$ penny flipover [Meyer and Schmidt (2000)], quantum Monty Hall problem [Flitney and Abbott (2002)], and quantum truel [Flitney and Abbott (2004)].

The protocol of the quantum sequential games is formulated as follows.

A quantum sequential game is a game $(T,R)$ which has the following form.

(1) $T = (N, \mathcal{H}, \hat{\rho}_{\text{init}}, Q, \mu, \{\hat{M}_r\}, \{a_r\})$, and each component satisfies the following conditions: (i) $N = \{1,2,\ldots,n\}$ is a finite set; (ii) $\mathcal{H}$ is a Hilbert space; (iii) $\hat{\rho}_{\text{init}}$ is a density operator on $\mathcal{H}$; (iv) $Q = Q_1 \times Q_2 \times \cdots \times Q_m$, where $Q_k$ is a subset of the set of all CPTP maps on the set of density operators on $\mathcal{H}$; the total number of operations is denoted by $m$; (v) $\mu$ is a bijection from $\bigcup_{i=1}^{n} \{(i,j)|1 \leq j \leq m_i\}$ to $\{1,\ldots,m\}$, where $m_i$’s are natural numbers satisfying $m_1 + \cdots + m_n = m$. The meaning of $\mu$ is that the $j$th operation for player $i$ is the $\mu(i,j)$-th operation in total; (vi) $\{\hat{M}_r\}$ is a POVM on $\mathcal{H}$; (vii) $a_r = (a_{r1}, a_{r2}, \ldots, a_{rn}) \in \mathbb{R}^n$; the index $r$ of $a_r$ runs over the same domain as that of $\hat{M}_r$.

(2) $R(T) = (N, \Omega, f)$. The strategy space $\Omega = \Omega_1 \times \cdots \times \Omega_n$ is constructed as

$$\Omega_i = Q_{\mu(i,1)} \times Q_{\mu(i,2)} \times \cdots \times Q_{\mu(i,m_i)}.$$
The value of $f$ is given by

$$f \left( (\mathcal{E}_{\mu(1)}, \ldots, \mathcal{E}_{\mu(1,m_1)}), \ldots, (\mathcal{E}_{\mu(n,1)}, \ldots, \mathcal{E}_{\mu(n,m_n)}) \right) = \sum_r \alpha_r \text{Tr} \left[ \hat{M}_r \mathcal{E}_m \circ \mathcal{E}_{m-1} \circ \cdots \circ \mathcal{E}_1 (\hat{\rho}_{\text{init}}) \right]$$

for all $((\mathcal{E}_{\mu(1)}, \ldots, \mathcal{E}_{\mu(1,m_1)}), \ldots, (\mathcal{E}_{\mu(n,1)}, \ldots, \mathcal{E}_{\mu(n,m_n)}) \in \Omega$.

If $\mathcal{H}$ is finite dimensional, then we refer to the game $(T, R)$ as a finite quantum sequential game. We denote the set of all quantum sequential games by $Q_{\text{Seq}}$, and the set of all finite quantum sequential games by $FQ_{\text{Seq}}$.

### 6.3.3 Equivalence of Formal Games

Now we define equivalence between two games. The basic idea is that two games are equivalent if their normal forms have the same structure, for the essence of a game is a game situation which is modelled by a normal form. The difficulty of this idea is that a strategy set $\Omega_i$ may have some redundancy; that is, two or more elements in $\Omega_i$ may represent essentially the same strategy. If this is the case, it does not work well to compare the strategy sets directly to judge whether two games are equivalent or not. Instead, we should define a new normal form in which the redundancy in the strategy set is excluded from the original normal form, and then compare the new normal forms of the two games.

As the first step to define equivalence between games, we clarify what it means by "two elements in $\Omega_i$ represent essentially the same strategy".

Let $(N, \Omega, f)$ be a normal form of a game. Two strategies $\omega_i, \omega'_i \in \Omega_i$ for player $i$ are said to be redundant if

$$f(\omega_1 \ldots \omega_{i-1}, \omega_i, \omega_{i+1} \ldots \omega_n) = f(\omega_1 \ldots \omega_{i-1}, \omega'_i, \omega_{i+1} \ldots \omega_n)$$

for all $\omega_1 \in \Omega_1, \ldots, \omega_{i-1} \in \Omega_{i-1}, \omega_{i+1} \in \Omega_{i+1}, \ldots, \omega_n \in \Omega_n$. If two strategies $\omega_i, \omega'_i \in \Omega_i$ are redundant, we write $\omega_i \sim \omega'_i$.

We can show that the binary relation $\sim$ is an equivalence relation. Namely, for all elements $\omega, \omega', \omega''$ of $\Omega_i$, the following holds: (i) $\omega \sim \omega'$; (ii) If $\omega \sim \omega'$ then $\omega' \sim \omega$; and (iii) If $\omega \sim \omega'$ and $\omega' \sim \omega''$ then $\omega \sim \omega''$.

Since $\sim$ is an equivalence relation, we can define the quotient set $\bar{\Omega}_i$ of a strategy set $\Omega_i$ by $\sim$. The quotient set $\bar{\Omega}_i$ is the set of all equivalence
classes in $\Omega_i$. An equivalence class in $\Omega_i$ is a subset of $\Omega_i$ which has the form of $\{ \omega | \omega \in \Omega_i, a \sim \omega \}$, where $a$ is an element of $\Omega_i$. We denote by $[\omega]$ an equivalence class in which $\omega$ is included, and we define $\tilde{\Omega}$ as $\tilde{\Omega} \equiv \tilde{\Omega}_1 \times \ldots \times \tilde{\Omega}_n$. This $\tilde{\Omega}$ is a new strategy set which has no redundancy.

Next, we define a new expected pay–off function $\tilde{f}$ which maps $\tilde{\Omega}$ to $\mathbb{R}^n$ by

$$\tilde{f}([\omega_1], \ldots, [\omega_n]) = f(\omega_1, \ldots, \omega_n).$$

This definition says that for $(C_1, \ldots, C_n) \in \tilde{\Omega}$, the value of $\tilde{f}(C_1, \ldots, C_n)$ is determined by taking one element $\omega_i$ from each $C_i$ and evaluating $f(\omega_1, \ldots, \omega_n)$. $\tilde{f}$ is well-defined. That is to say, the value of $\tilde{f}(C_1, \ldots, C_n)$ is independent of which element in $C_i$ one would choose. To show this, suppose $(C_1, \ldots, C_n) \in \tilde{\Omega}$ and $\alpha_i, \beta_i \in C_i$. Then $\alpha_i \sim \beta_i$ for every $i$, so that

$$f(\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n) = f(\beta_1, \alpha_2, \alpha_3, \ldots, \alpha_n)$$

$$= f(\beta_1, \beta_2, \alpha_3, \ldots, \alpha_n)$$

$$\vdots$$

$$= f(\beta_1, \beta_2, \beta_3, \ldots, \beta_n).$$

Thus the value of $\tilde{f}(C_1, \ldots, C_n)$ is determined uniquely.

Using $\tilde{\Omega}$ and $\tilde{f}$ constructed from the original normal form $(N, \Omega, f)$, we define the new normal form as follows.

Let $(N, \Omega, f)$ be the normal form of a game $G$. We refer to $(N, \tilde{\Omega}, \tilde{f})$ as the reduced normal form of $G$.

Whether two games are equivalent or not is judged by comparing the reduced normal forms of these games, as we mentioned earlier.

Let $(N^{(1)}, \tilde{\Omega}^{(1)}, \tilde{f}^{(1)})$ be the reduced normal form of a game $G_1$, and let $(N^{(2)}, \tilde{\Omega}^{(2)}, \tilde{f}^{(2)})$ be the reduced normal form of a game $G_2$. Then, $G_1$ is said to be equivalent to $G_2$ if the following holds: (i) $N^{(1)} = N^{(2)} = \{1, \ldots, n\}$; and (ii) There exists a sequence $(\phi_1, \ldots, \phi_n)$ of bijection $\phi_k : \tilde{\Omega}_k^{(1)} \mapsto \tilde{\Omega}_k^{(2)}$, such that for all $(C_1, \ldots, C_n) \in \tilde{\Omega}^{(1)}$

$$\tilde{f}^{(1)}(C_1, \ldots, C_n) = \tilde{f}^{(2)}(\phi_1(C_1), \ldots, \phi_n(C_n)).$$

(6.43)

If $G_1$ is equivalent to $G_2$, we write $G_1 \parallel G_2$.

To give an example of equivalent games, let us consider classical $PQ$ penny flipover [Meyer and Schmidt (2000)], in which both player $P$ and player $Q$ are classical players. In this game, a penny is placed initially.
heads up in a box. Players take turns \((Q \rightarrow P \rightarrow Q)\) flipping the penny over or not. Each player can not know what the opponent did, nor see inside the box. Finally the box is opened, and \(Q\) wins if the penny is heads up. This game can be formulated as a finite strategic game whose pay–off matrix is given in Table 6.6.

Table 6.6 Pay–off matrix for \(PQ\) penny flipover. \(F\) denotes a flipover and \(N\) denotes no flipover. The first entry in the parenthesis denotes \(P\)’s pay–off and the second one denotes \(Q\)’s pay–off.

<table>
<thead>
<tr>
<th></th>
<th>(Q: NN)</th>
<th>(Q: NF)</th>
<th>(Q: FN)</th>
<th>(Q: FF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P: N)</td>
<td>((-1, 1))</td>
<td>((1, -1))</td>
<td>((1, -1))</td>
<td>((-1, 1))</td>
</tr>
<tr>
<td>(P: F)</td>
<td>((1, -1))</td>
<td>((-1, 1))</td>
<td>((-1, 1))</td>
<td>((1, -1))</td>
</tr>
</tbody>
</table>

Intuitively, \(Q\) does not benefit from the second move, so that it does not matter whether \(Q\) can do the second move or not. The notion of equivalence captures this intuition; the above penny flipover game is equivalent to a finite strategic game whose pay–off matrix is given in Table 6.7. It represents another penny flipover game in which both players act only once. Proof of the equivalence is easy and we omit it.

Table 6.7 Pay–off matrix for another \(PQ\) penny flipover in which both players act only once.

<table>
<thead>
<tr>
<th></th>
<th>(Q: N)</th>
<th>(Q: F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P: N)</td>
<td>((-1, 1))</td>
<td>((1, -1))</td>
</tr>
<tr>
<td>(P: F)</td>
<td>((1, -1))</td>
<td>((-1, 1))</td>
</tr>
</tbody>
</table>

We now return to the general discussion on the notion of equivalence. The following is a basic property of the equivalence between two games. The binary relation \(\parallel\) is an equivalence relation; namely, for any games \(G_1, G_2,\) and \(G_3,\) the following holds: reflexivity: \(G_1 \parallel G_1;\) symmetry: If \(G_1 \parallel G_2,\) then \(G_2 \parallel G;\) and transitivity: If \(G_1 \parallel G_2\) and \(G_2 \parallel G_3,\) then \(G_1 \parallel G_3.\) For the proof, see [Kobayashi (2007)].

In some cases, we can find that two games are equivalent by comparing the normal forms of the games, not the reduced normal forms. In the following lemma, sufficient conditions for such cases are presented.

Let \((N^{(1)}, \Omega^{(1)}, f^{(1)})\) be the normal form of a game \(G_1,\) and let \((N^{(2)}, \Omega^{(2)}, f^{(2)})\) be the normal form of a game \(G_2.\) If the following conditions are satisfied, \(G_1\) is equivalent to \(G_2:\) (i) \(N^{(1)} = N^{(2)} = \{1, \ldots, n\};\) and (ii) There exists a sequence \((\psi_1, \ldots, \psi_n)\) of bijection \(\psi_k : \Omega^{(1)}_k \rightarrow \Omega^{(2)}_k,\) such that for all \((\omega_1, \ldots, \omega_n) \in \Omega^{(1)},\)

\[
f^{(1)}(\omega_1, \ldots, \omega_n) = f^{(2)}(\psi_1(\omega_1), \ldots, \psi_n(\omega_n)).
\] (6.44)
For the proof, see [Kobayashi (2007)].

We define a map \( \phi_i \) from \( \Omega^{(1)}_i \) to the set of all subsets of \( \Omega^{(2)}_i \) as

\[
\phi_i(C_i) = \{ \psi_i(\omega') | \omega' \in C_i \}.
\]

We show that the range of \( \phi_i \) is a subset of \( \tilde{\Omega}^{(2)}_i \); that is, for any \( [\omega_i] \in \tilde{\Omega}^{(1)}_i \) there exists \( \xi_i \in \Omega^{(2)}_i \) such that \( \phi_i([\omega_i]) = [\xi_i] \). In fact, \( \psi_i(\omega_i) \) is such a \( \xi_i \):

\[
\phi_i([\omega_i]) = [\psi_i(\omega_i)]. \tag{6.45}
\]

Below, we will prove \( \phi_i([\omega_i]) \subset [\psi_i(\omega_i)] \) first, and then prove \( [\psi_i(\omega_i)] \subset \phi_i([\omega_i]) \).

To prove \( \phi_i([\omega_i]) \subset [\psi_i(\omega_i)] \), we will show that an arbitrary element \( \sigma_i \in \phi_i([\omega_i]) \) satisfies \( \sigma_i \in [\psi_i(\omega_i)] \). For this purpose, it is sufficient to show that \( \sigma_i \sim \psi_i(\omega_i) \); that is, for an arbitrary \( \sigma_k \in \Omega^{(2)}_k \) \((k \neq i)\)

\[
f^{(2)}(\sigma_1, \ldots, \sigma_{i-1}, \sigma_i, \sigma_{i+1}, \ldots, \sigma_n) = f^{(2)}(\sigma_1, \ldots, \sigma_{i-1}, \psi_i(\omega_i), \sigma_{i+1}, \ldots, \sigma_n). \tag{6.46}
\]

Since \( \psi_k \) is a bijection, there exists \( \omega_k \in \Omega^{(1)}_k \) such that \( \psi_k(\omega_k) = \sigma_k \). In addition, because \( \sigma_i \in \phi_i([\omega_i]) \), there exists \( \omega_i' \in [\omega_i] \) such that \( \sigma_i = \psi_i(\omega_i') \). Thus,

\[
f^{(2)}(\sigma_1, \ldots, \sigma_{i-1}, \sigma_i, \sigma_{i+1}, \ldots, \sigma_n)
= f^{(2)}(\psi_i(\omega_i), \ldots, \psi_{i-1}(\omega_{i-1}), \psi_i(\omega_i'), \ldots, \psi_n(\omega_n))
= f^{(1)}(\omega_i, \omega_{i-1}, \omega_i', \omega_{i+1}, \ldots, \omega_n). \tag{6.47}
\]

Because \( \omega_i' \in [\omega_i] \) and \( \omega_i \in [\omega_i] \), it follows that \( \omega_i' \sim \omega_i \). Hence,

\[
\text{(6.47)} = f^{(1)}(\omega_i, \omega_{i-1}, \omega_i', \omega_{i+1}, \ldots, \omega_n)
= f^{(2)}(\psi_i(\omega_i), \ldots, \psi_{i-1}(\omega_{i-1}), \psi_i(\omega_i'), \ldots, \psi_n(\omega_n))
= f^{(2)}(\sigma_1, \ldots, \sigma_{i-1}, \psi_i(\omega_i), \sigma_{i+1}, \ldots, \sigma_n), \tag{6.48}
\]

which leads to the conclusion that the equation (6.46) holds for any \( \sigma_i \in \phi_i([\omega_i]) \).

Conversely, we can show that \( [\psi_i(\omega_i)] \subset \phi_i([\omega_i]) \). Let \( \sigma_i \) be an arbitrary element of \( [\psi_i(\omega_i)] \). Since \( \psi_i \) is a bijection, there exists \( \omega_i' \in \Omega^{(1)}_i \) such that \( \psi_i(\omega_i') = \sigma_i \). For such \( \omega_i' \), it holds that \( \psi_i(\omega_i') \sim \psi_i(\omega_i) \), because
ψ_i(ω'_j) ∈ [ψ_i(ω_i)]. Hence,

\[
f^{(1)}(ω_1, \ldots, ω_{i-1}, ω'_i, ω_{i+1}, \ldots, ω_n) = f^{(2)}(ψ_1(ω_1), \ldots, ψ_{i-1}(ω_{i-1}), ψ_i(ω'_i), ψ_{i+1}(ω_{i+1}), \ldots, ψ_n(ω_n)) = f^{(2)}(ψ_1(ω_1), \ldots, ψ_{i-1}(ω_{i-1}), ψ_i(ω_i), ψ_{i+1}(ω_{i+1}), \ldots, ψ_n(ω_n)) = f^{(1)}(ω_1, \ldots, ω_{i-1}, ω_i, ω_{i+1}, \ldots, ω_n),
\]
which indicates that ω'_j ∼ ω_i. Thus, ω'_j ∈ [ω_i]. Therefore, we conclude that if σ_i ∈ [ψ_i(ω_i)], then σ_i = ψ_i(ω'_i) ∈ φ_i([ω_i]); that is, [ψ_i(ω_i)] ⊂ φ_i([ω_i]).

We have shown above that φ_i is a map from ˜Ω_i^{(1)} to ˜Ω_i^{(2)}. The next thing we have to show is that φ_i is a bijection from ˜Ω_i^{(1)} to ˜Ω_i^{(2)}. We will show the bijectivity of φ_i by proving injectivity and surjectivity separately.

First, we show that φ_i is injective. Suppose [ω_i], [ω'_i] ∈ ˜Ω_i^{(1)} and [ω_i] ≠ [ω'_i]. Because [ω_i] ≠ [ω'_i], it follows that ω_i ∼ ω'_i, so that there exists (ω_1, ω_2, ω_3, ω_4, ω_5, ω_6, ω_7) ∈ ˜Ω_i^{(1)} × ˜Ω_i^{(1)} × ˜Ω_i^{(1)} × ˜Ω_i^{(1)} × ˜Ω_i^{(1)} × ˜Ω_i^{(1)} × ˜Ω_i^{(1)} such that

\[
f^{(1)}(ω_1, \ldots, ω_{i-1}, ω_i, ω_{i+1}, \ldots, ω_n) ≠ f^{(1)}(ω_1, \ldots, ω_{i-1}, ω'_i, ω_{i+1}, \ldots, ω_n).
\]

For such (ω_1, ω_2, ω_3, ω_4, ω_5, ω_6, ω_7),

\[
f^{(2)}(ψ_1(ω_1), \ldots, ψ_{i-1}(ω_{i-1}), ψ_i(ω_i), ψ_{i+1}(ω_{i+1}), \ldots, ψ_n(ω_n)) = f^{(1)}(ω_1, \ldots, ω_{i-1}, ω_i, ω_{i+1}, \ldots, ω_n) ≠ f^{(1)}(ω_1, \ldots, ω_{i-1}, ω'_i, ω_{i+1}, \ldots, ω_n) = f^{(2)}(ψ_1(ω_1), \ldots, ψ_{i-1}(ω_{i-1}), ψ_i(ω'_i), ψ_{i+1}(ω_{i+1}), \ldots, ψ_n(ω_n)).
\]

This indicates that ψ_i(ω_i) ∼ ψ_i(ω'_i). Hence, [ψ_i(ω_i)] ≠ [ψ_i(ω'_i)]. Thus, using (6.45), we conclude that φ_i([ω_i]) ≠ φ_i([ω'_i]).

Next, we show that φ_i is surjective. Let [σ] be an arbitrary element of ˜Ω_i^{(2)}. Define ω ∈ ˜Ω_i^{(1)} as ω ≡ ψ_i^{-1}(σ). Then,

\[
φ_i([ω]) = [ψ_i(ω)] = [σ],
\]

The first equation follows from (6.45). Thus, for an arbitrary [σ] ∈ ˜Ω_i^{(2)}, there exists [ω] ∈ ˜Ω_i^{(1)} such that φ_i([ω]) = [σ].
Lastly, we show that \((\varphi_1, \ldots, \varphi_n)\) satisfies (6.43). For an arbitrary
\(([\omega_1], \ldots, [\omega_n]) \in \tilde{\Omega}^{(1)}\),
\[
\tilde{f}^{(1)}([\omega_1], \ldots, [\omega_n]) = f^{(1)}(\omega_1, \ldots, \omega_n) = f^{(2)}(\psi_1(\omega_1), \ldots, \psi_n(\omega_n)) = \tilde{f}^{(2)}(\phi_1([\omega_1]), \ldots, \phi_n([\omega_n])).
\]
Equations (6.49) and (6.51) follow from the definition of \(\tilde{f}^{(1)}\) and \(\tilde{f}^{(2)}\).
Equation (6.50) follows from (6.44). The last equation follows from (6.45).

6.3.4 Game Classes

This short subsection is devoted to explaining game classes and some binary
relations between game classes. These notions simplify the statements of
our main theorems.

First, we define a game class as a subset of \(G\). We defined previously
\(G, SG, FSG, MEFSG, QSim, FQSim, QSeq, \) and \(FQSeq\). All of these are
game classes. Note that \(G\) is itself a game class.

Next, we introduce some symbols. Let \(A\) and \(B\) be game classes. If for
any game \(G \in A\) there exists a game \(G' \in B\) such that \(G \parallel G'\), then we
write \(A \triangleright B\). If \(A \triangleright B\) and \(B \triangleright A\), we write \(A \bowtie B\).
Equations (6.49) and (6.51) follow from the definition of \(\tilde{f}^{(1)}\) and \(\tilde{f}^{(2)}\).

6.3.5 Summary on Formal Quantum Games

The following relationships exist between game classes \(QSim, QSeq, FQSim,\) and \(FQSeq\) [Kobayashi (2007)]:
\(QSim \sqsubseteq QSeq\). Also, we have:
\(FQSim \sqsubseteq FQSeq\) and the converse \(FQSeq \sqsubseteq FQSim\). For the proof, see [Kobayashi (2007)].

When a statement “if \(G \in FQSim\) then \(G\) has a property \(P\)” is true,
another statement “if \(G \in FQSeq\) then \(G\) has a property \(P\)” is also true,
and vice versa. Here, \(P\) must be such a property that if a game \(G\) has the
property \(P\) and \(G \parallel G'\), then \(G'\) also has the property \(P\). We call such \(P\)
a property preserved under \(\parallel\). For example, “a Nash equilibrium exists” is
a property preserved under \(\parallel\).
Fig. 6.5 A quantum sequential game $G^{seq}$ which is equivalent to a quantum simultaneous game $G$ depicted in Figure 6.3 (see text for explanation).

Fig. 6.6 A finite quantum simultaneous game $G^{sim}$ which is equivalent to a given finite quantum sequential game $G$ (see text for explanation).

Unfortunately, no results are known which have the form “if $G \in FQSim (FQSeq)$ then $G$ has a property $P$, but otherwise $G$ does not necessarily have the property $P$”. Consequently, we cannot reap the benefits of the above-mentioned deduction. However, numerous results exist which have the form “for a certain subset $S$ of $FQSim (FQSeq)$, if $G \in S$ then $G$ has a property $Q$ preserved under $\parallel$, but otherwise $G$ does not necessarily have the property $Q$”. For such $S$ and $Q$, the above theorems guarantee that there exists a subset $S'$ of $FQSeq (FQSim)$ which satisfies the following: “If $G \in S'$ then $G$ has the property $Q$, but otherwise $G$ does not necessarily have the property $Q$”. In this sense, many of the results so far on $FQSim (FQSeq)$ can be translated into statements on $FQSeq (FQSim)$ [Kobayashi (2007)].
It is worth noting that efficiency of a game is not a property preserved under $\|\|$. In an original quantum sequential game $G$, it is necessary to transmit a qudit $m + 1$ times, while $4m$ times are needed in the constructed game $G^{\text{sim}}$. Thus, $G^{\text{sim}}$ is far more inefficient than $G$, despite $G^{\text{sim}}$ and $G$ are equivalent games.

Relevant to the present subsection is the study by Lee and Johnson\cite{Lee and Johnson (2003)}. To describe their argument, we have to introduce a new game class.

A finite quantum simultaneous game with all CPTP maps available is a subclass of finite quantum simultaneous games, in which a strategy set $\Omega_i$ is the set of all CPTP maps on the set of density operators on $H_i$ for every $i$. We denote the set of all finite quantum simultaneous game with all CPTP maps available by $\text{FQSimAll}$.

We can easily prove that $\text{FQSimAll} \triangleleft \text{FQSim}$, by showing that the range of expected pay–off functions for a game in $\text{FQSimAll}$ must be connected, while the one for a game in $\text{FQSim}$ can be disconnected.

Lee and Johnson claimed that (i) “any game could be played classically” and (ii) “finite classical games consist of a strict subset of finite quantum games”. Using the terms of this subsection, we may interpret these claims as follows.

We have: $\text{SG} \bowtie G$, $\text{MEFG} \preceq \text{FQSimAll}$ and $\text{FQSimAll} \not\preceq \text{MEFG}$. For the proof, see \cite{Kobayashi (2007)}.

The relationships between various game classes can be summarized as:

\[
\text{MEFG} \triangleleft \text{FQSimAll} \triangleleft \text{FQSim} \bowtie \text{FQSeq}
\]
\[
\preceq \text{QSim} \preceq \text{QSeq} \preceq \text{SG} \bowtie G.
\] (6.53)

Replacing $\preceq$ in (6.53) with either $\triangleleft$ or $\bowtie$ will be a possible extension of this subsection.

Besides that, it remains to be investigated what the characterizing features of each game class in (6.53) are. Especially, further research on games which is in $\text{FQSim}$ (or equivalently $\text{FQSeq}$) but not equivalent to any games in $\text{MEFG}$ would clarify the truly quantum mechanical nature of quantum games.

19 Amount of information exchange between players and a referee, required to play a game. See [Lee and Johnson (2003)] for more details.
Quantum Games and Quantum Computers

6.4 Quantum Information and Computing

6.4.1 Entanglement, Teleportation and Information

Any storage, transmission, and processing of information relies on a physical carrier [Landauer (1991)]. In a handwritten note the sheet of paper serves as the carrier of information, in a desk top computer it is the random access memory and the hard drive on which the relevant data are stored. Communication makes use of sound waves, radio waves, or light pulses. The new field of quantum information is based on the idea that single quantum systems can be used as the elementary carriers of information, such as single photons, atoms, and ions. Quantum theory opens up new possibilities for information processing and communication [Bouwmeester et al. (2000); Ekert and Jozsa (1996); Nielsen and Chuang (2000); Steane (1998); Ekert and Jozsa (1996); Plenio and Vedral (1998a)]. In principle, quantum systems allow for information processing tasks which are very difficult or impossible to be performed when using classical systems. Envisioned applications range from the factorization of large numbers on a quantum computer to communication protocols, and key distribution in quantum cryptography.

Quantum theory may become relevant to technical development in information processing mainly for two reasons [Eisert (2001)]. On the one hand, the information processing and storage units in ordinary, “classical” computers are becoming smaller and smaller. The dimensions of transistor elements in silicon-based microchips are decreasing to the extent that they will be approaching scales in which quantum effects become relevant in the near future. On the other hand, it has become technologically possible to store and manipulate single quantum systems, e.g., with sophisticated methods from quantum optics and solid state physics [Bouwmeester et al. (2000); Nielsen and Chuang (2000)].

The superior “performance” of quantum systems in computation and communication applications is predominantly rooted in a property of quantum mechanical states called entanglement. Essentially, entanglement comes along with new kinds of correlations. Even in classical composite systems measurement outcomes may be perfectly correlated. However, entangled quantum states may show stronger statistical correlations than those attainable in a classical composite system, where the correlation is produced by a classical random generator.\(^{20}\)

\(^{20}\)In 1935 A. Einstein, B. Podolsky, and N. Rosen (EPR) published a seminal paper...
On a purely theoretical level it is important to understand what kinds of tasks may be achieved with entangled quantum systems. For example, it is impossible to transmit the particular “quantum information” of a quantum system through a classical channel. This means that the statistical predictions of quantum mechanics cannot fully be reproduced when trying to extract information about the preparation procedure from a state of one quantum system, transmitting the classical information through the channel, and preparing another quantum system in a certain state on the basis of this information. It is nevertheless possible to transfer a quantum state to a different quantum system at a distant location without physically transmitting the actual quantum system in the particular state – provided that the parties initially share a pair of two–level systems in a maximally entangled state. This transfer of a quantum state was named quantum teleportation, a term borrowed from science–fiction literature [Eisert (2001)].

The teleportation protocol (see Figure 6.7) was proposed in 1993 by C. Bennett et al. [Bennett et al. (1993)]. It represented a major breakthrough in the field. Assume that one experimenter, from now on referred to as Alice, would like to send the unknown state of a given quantum system to Bob, another experimenter at a distant location. Bob prepares a bi–partite quantum system in a particular entangled state and gives one part of the system to Alice. In the next step Alice performs a local joint quantum measurement on both her quantum system and on the one she has received from Bob. Then she phones Bob and tells him about the outcome. Depending on the outcome of Alice’s measurement, Bob can finally transform

entitled “Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?”, which started a long lasting debate about the status of quantum theory [Einstein et al. (1935b)]. On the basis of the predicted outcomes of measurements on two space–like separated quantum particles in an entangled state, EPR came to the conclusion that quantum mechanics could not be a complete theory, suggesting the view that additional hidden variables should be appended to a quantum state in order to restore causality and locality. N. Bohr, one of the spokesmen of the so–called Copenhagen school of the interpretation of quantum mechanics, argued against the assumption of a more complete underlying local deterministic level [Bohr (1935)]. It was not until 1964 that J. Bell presented a way to empirically test the two competing hypotheses [Bell (1964); Bell (1966); Bell (1987)]. Bell’s theorem is not strictly about quantum mechanics. It is a statement concerning correlations of measurement outcomes at distant events that any physical theory may predict under the assumption of an underlying local classical model [Mermin (1993); Ballentine (1987)]. Starting from the 1980s many experiments were performed, each in favor of quantum mechanics and against local hidden variable models [Aspect et al. (1981); Clauser and Shimony (1990); Tittel et al. (1998)]. More on EPR will be given in subsection 7.1.2 below.
Fig. 6.7 The teleportation protocol as in [Bennett et al. (1993)] (modified and adapted from [Eisert (2001)]).

his part of the maximally entangled system by use of a simple manipulation of the state. The state of his system is eventually identical to the state of Alice’s original system: the state has been “sent” from Alice to Bob.

Another important proposal of this type, also by C. Bennett and S. Wiesner, is the dense coding protocol [Bennett and Wiesner (1992)] concerning the transmission of classical information. A single quantum two–level system sent from Alice to Bob can carry a single bit of classical information. Surprisingly, if the two parties initially share a maximally entangled state, two bits of classical information can be transmitted with the same two–level system. Successful experimental quantum optical implementations of dense coding and teleportation were performed by [Bouwmeester et al. (1997)] and [Boschi et al. (1998)]. Following the publication of these proposals, many other applications of entanglement were suggested [21], the spectrum ranging from quantum cryptography using

However, one has to bear in mind that notwithstanding the possible technical implications the research in quantum entanglement is to a large extent motivated by the wish to better understand one of the most important properties that distinguishes quantum mechanics from classical mechanics. Recent research in the theory of entanglement shows that there are still fundamental open questions in non-relativistic quantum mechanics waiting to be solved.
entangled quantum systems by [Ekert (1991)] to improved frequency standards [Cirac et al. (1998)] and clock synchronization [Jozsa et al. (2000)].

6.4.2 Quantum Computing

It is well known that we do not need quantum mechanics in order to explain how the bits, that is the zeros and ones, inside a classical computer evolve. The reason for this is that the architecture of classical computers does not make use of one of the most fundamental features of quantum mechanics, namely the possibility of superpositions. Throughout the entire processing of any program on a classical computer, each of the involved bits takes on either the value zero or one. Quantum mechanics, however, would in addition allow superpositions of zeros and ones, that is, quantum–bits, or qubits, which are somehow in the state zero and one at the same time. Computing devices which exploit this possibility, and with that all the essential features of quantum mechanics, are popularly called quantum computers [Nielsen and Chuang (2000)]. Since they have an additional capability they are at least as powerful as classical computers: every problem that can be solved on a classical computer can be handled by a quantum computer just as well. The converse, however, is also true since the dynamics of quantum systems is governed by linear differential equations, which can in turn be solved (at least approximately) on a classical computer. Hence [Eisert and Wolf (2006)], classical and quantum computers could in principle emulate each other and quantum computers are thus no hypercomputers. So why quantum computing? And if there is any reason, why not just simulate these devices (which do not exist yet anyhow) on a classical computer?

One reason for aiming at building quantum computers is that they will solve certain types of problems faster than any (present or future) classical computer – it seems that the border between easy and hard problems is different for quantum computers than it is for their classical counterparts [Eisert and Wolf (2006)]. Here ‘easy’ means that the time for solving the problem grows polynomially with the length of the input data (like for the problem of multiplying two numbers), whereas ‘hard’ problems are those for which the required time grows exponentially. Prominent examples for

22A hypercomputer would be capable of solving problems that can not be handled by a universal Turing machine (the paradigm of a classical digital computer). The most famous example of such a problem is the halting problem which is in modern terminology the task of a universal crash debugger, which is supposed to spot all bugs leading to crashes or infinite loops for any program running on a universal Turing machine. As shown by Turing such a debugger cannot exist.
hard problems are the travelling salesman problem, the graph isomorphism problem, and the problem of factoring a number into primes.\textsuperscript{23} For the latter it was, to the surprise of all, shown by Shor in 1994 that it could efficiently be solved by a quantum computer in polynomial time \cite{Shor_1997}. Hence, a problem which is hard for any classical computer becomes easy for quantum computers.\textsuperscript{24} Shor’s result gets even more brisance from the fact that the security of public key encryption, i.e., the security of home banking and any other information transfer via the internet, is heavily based on the fact that factoring is a hard problem.

On the other hand, nature provides many fascinating collective quantum phenomena like superconductivity, magnetism and Bose–Einstein condensation. Although all properties of matter are described by and can in principle be determined from the laws of quantum mechanics, physicists have very often serious difficulties to understand them in detail and to predict them by starting from fundamental rules and first principles. One reason for these difficulties is the fact that the number of parameters needed to describe a many-particle quantum system grows exponentially with the number of particles. Hence, comparing a theoretical model for the behavior of more than, say, thirty particles with experimental reality is not possible by simulating the theoretical model numerically on a classical computer without making serious simplifications \cite{Eisert_Wolf_2006}.

When thinking about this problem of simulating quantum systems on classical computers Richard Feynman came in the early eighties to the conclusion that such a classical simulation typically suffers from an exponential slowdown, whereas another quantum system could in principle do the simulation efficiently with bearable overhead \cite{Feynman_1996}. In this way a quantum computer operated as a quantum simulator could be used as a link between theoretical models which are formulated on a fundamental level and experimental observations. Similar to Shor’s algorithm, a quantum simulator would yield an exponential speedup compared to a classical computer. An important difference between these two applications is, how-

\textsuperscript{23}These problems are strongly believed to be hard (the same is by the way true for a special instance of the computer game “minesweeper”). However, in all cases there is no proof that a polynomial-time algorithm can not exist. The question whether there exists such an algorithm (for the travelling salesman or the minesweeper problem) is in fact the notorious $P \overset{?}{=} NP$ question for whose solution there is even a prize of one million dollar.

\textsuperscript{24}In fact Shor’s algorithm strikes the strong Church–Turing thesis, which states that every reasonable physical computing device can be simulated on a probabilistic Turing machine with at most polynomial overhead.
ever, that a useful Shor–algorithm quantum computer requires thousands of qubits whereas a few tens of qubits could already be useful for the simulation of quantum systems.

Apart from the computational power of a quantum computer, there is a much more banal argument for incorporating quantum mechanics into computer science: Moore’s law. In 1965 Intel co-founder G. Moore observed an exponential growth in the number of transistors per square inch on integrated circuit and he predicted that this trend would continue [Moore (1989)]. In fact, since then this density has doubled approximately every 18 months. If this trend continues then around the year 2020 the components of computers are at the atomic scale where quantum effects are dominant. We have thus to inevitably cope with these effects, and we can either try to circumvent and eliminate them as long as this is possible and keep on doing classical computing, or we can at some point try to make use of them and start doing quantum computing [Eisert and Wolf (2006)].

Besides the quantum computer with its mentioned applications, quantum information science yields a couple of other useful applications which might be easier to realize. The best example is quantum cryptography which enables one to transmit information with “the security of nature’s laws” [Gisin et al. (2002)]. However, small building blocks of a quantum computer, i.e., small quantum circuits may be useful as well. One potential application is for instance in precision measurements like in atomic clocks [Wineland et al. (1992); Huelga et al. (1997)]. The latter are important in global positioning systems as well as in synchronizing networks and distant telescopes. By generating quantum correlations between the $N$ relevant atoms in the atomic clock, a quantum circuit could in principle reduce the uncertainty of the clock by a factor $\sqrt{N}$.

Another application of small quantum circuits is entanglement distillation: in order to distribute entangled states over large distances we have to send them through inevitably noisy channels, thereby loosing some of the entanglement. Fortunately, however, we can in many cases distill a few highly entangled states out of many weakly entangled ones [Bennett et al. (1996); Horodecki and Horodecki (2001)].

Let us now compare the concepts of classical computing with the basics of quantum computing. In fact, many classical concepts have very similar quantum counterparts, like bits become qubits and still the logic is often best explained within a circuit model [Deutsch (1989); Nielsen and Chuang (2000)]. However, there are also crucial differences.
6.4.2.1 Qubits and Quantum Parallelism

Recall that the elementary information carriers in a quantum computer are the quantum–bits, or qubits [Schumacher (1995)]. In contrast to classical bits which take on either the value zero or one, qubits can be in every superposition of the state vectors $|0\rangle$ and $|1\rangle$. This means that the vector $|\Psi\rangle$ describing the pure state of the qubit can be any linear combination

$$|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

(6.54)

of the vectors $|0\rangle$ and $|1\rangle$ with complex coefficients $\alpha$ and $\beta$. In the same way a system of many qubits can be in a superposition of all classically possible states

$$|0,0,\ldots,0\rangle + |1,0,\ldots,0\rangle + \ldots + |1,1,\ldots,1\rangle.$$  

(6.55)

The basis $\{|0,0,\ldots,0\rangle, |0,1,\ldots,0\rangle, \ldots, |1,1,\ldots,1\rangle\}$ that corresponds to the binary words of length $n$ in a quantum system of $n$ qubits is called the computational basis.

Using the superposition of (6.55) as an input for an algorithm means somehow to run the computation on all classically possible input states at the same time. This possibility is called quantum parallelism and it is certainly one of the reasons for the computational power of a quantum computer. The mathematical structure behind the composition of quantum systems is the one of the tensor product. Hence, vectors like $|0,0,\ldots,0\rangle$ should be understood as $|0\rangle \otimes \ldots \otimes |0\rangle = |0\rangle^\otimes n$. This implies that the dimension of the space characterizing the system grows exponentially with the number of qubits [Eisert and Wolf (2006)].

Physically, qubits correspond to effective two–level systems like ground state and excited state of an atom, the polarization degree of freedom of light or up–and down orientation of a spin–1/2 particle. Such a physical system can be in any pure state that can be represented by a normalized vector of the above form. A pure state of a composite quantum system that is not a product with respect to all constituents is called an entangled pure state.

6.4.2.2 Quantum Computers

An important difference between classical and quantum computers lies in the read–out process. In the classical case there is not much to say: the output is a bit-string which is obtained in a deterministic manner, i.e., repeating the computation will lead to the same output again. However,
due to the probabilistic nature of quantum mechanics, this is different for a quantum computer. If the output of the computation is for instance the state vector $|\Psi\rangle$ in (6.54), $\alpha$ and $\beta$ cannot be determined by a single measurement on a single specimen. In fact, $|\alpha|^2$ and $|\beta|^2$ are the probabilities for the system to be found in $|0\rangle$ and $|1\rangle$ respectively. Hence, the absolute values of these coefficients can be determined by repeating the computation, measuring in the basis $|0\rangle$, $|1\rangle$ and then counting the relative frequencies. The actual outcome of every single measurement is thereby completely indetermined. In the same manner, the state of a quantum system consisting of $n$ qubits can be measured in the computational basis, which means that the outcome corresponding to some binary word occurs with the probability given by the square of the absolute value of the respective coefficient. So in effect, the probabilistic nature of the read out process on the one hand and the possibility of exploiting quantum parallelism on the other hand are competing aspects when it comes to comparing the computational power of quantum and classical computers [Eisert and Wolf (2006)].

6.4.2.3 The Circuit Model for Quantum Computers

Recall that a classical digital computer operates on a string of input bits and returns a string of output bits. The function in between can be described as a logical circuit build up out of many elementary logic operations. That is, the whole computation can be decomposed into an array of smaller operations, or gates, acting only on one or two bits like the AND, OR and NOT operation. In fact, these three gates together with the COPY (or FANOUT) operation form a universal set of gates into which every well-defined input–output function can be decomposed. The complexity of an algorithm is then essentially the number of required elementary gates, resp. its asymptotic growth with the size of the input [Eisert and Wolf (2006)].

The circuit model for the quantum computer [Deutsch (1989); Nielsen and Chuang (2000)] is actually very reminiscent of the classical circuit model: we only have to replace the input–output function by a quantum operation mapping quantum states onto quantum states. It is sufficient to consider operations only that have the property to be unitary, which means that the computation is taken to be logically reversible. In turn, any unitary operation can be decomposed into elementary gates acting only on one or two qubits. A set of elementary gates that allows for a realization of any unitary to arbitrary approximation is again referred to as being universal [Barenco et al. (1996)].
Deutsch (1989)]. An important example of a set of universal gates is in this case any randomly chosen one–qubit rotation together with the CNOT (Controlled NOT) operation, which acts as
\[ |x, y⟩ \mapsto |x, y \oplus x⟩, \tag{6.56} \]
where \( \oplus \) means addition modulo 2 [Preskill (1998)]. Like in the classical case there are infinitely many sets of universal gates. Notably, also any generic (i.e., randomly chosen) two–qubit gate (together with the possibility of switching the leads in order to swap qubits) is itself a universal set, very much like the NAND gate is for classical computing [Barenco et al. (1996)]. Notably, any quantum circuit that makes use of a certain universal set of quantum gates can be simulated by a different quantum circuit based on another universal set of gates with only poly–logarithmic overhead [Kitaev (1997); Nielsen and Chuang (2000)]. A particularly useful single-qubit gate is the Hadamard gate, acting as
\[ |0⟩ \mapsto H|0⟩ = (|0⟩ + |1⟩)/\sqrt{2}, \quad |1⟩ \mapsto H|1⟩ = (|0⟩ - |1⟩)/\sqrt{2}. \tag{6.57} \]
A **phase gate** does nothing but multiplying one of the basis vectors with a phase,
\[ |0⟩ \mapsto |0⟩, \quad |1⟩ \mapsto i|1⟩, \tag{6.58} \]
and a **Pauli gate** corresponds to one of the three unitary Pauli matrices. The CNOT, the Hadamard, the phase gate, and the Pauli gate are quantum gates of utmost importance. Given their key status in many quantum algorithms, one might be tempted to think that with these ingredients alone (together with measurements of Pauli operators, see below), powerful quantum algorithms may be constructed that outperform the best known classical algorithm to a problem. This intuition is yet not correct: it is the content of the Gottesman–Knill theorem that any quantum circuit consisting of only these ingredients can be simulated efficiently on a classical computer [Nielsen and Chuang (2000)].

\[ \text{Gottesman (1997).} \]

\[ \text{Any such generic quantum gate has so–called entangling power [Collins et al. (2001)], in that it may transform a product state vector into one that can no longer be written as a tensor product. Such quantum mechanical pure states are called entangled. In intermediate steps of a quantum algorithm the physical state of the system is in general highly multi–particle entangled. In turn, the implementation of quantum gates in distributed quantum computation requires entanglement as a resource [Eisert et al. (2000)].} \]

\[ 25 \]
One of the crucial differences between classical and quantum circuits is, that in the quantum case the COPY operation is not possible. In fact, the linearity of quantum mechanics forbids a device which copies an unknown quantum state – this is known as the no-cloning theorem. The latter has far-reaching consequences, of which the most prominent one is the possibility of quantum cryptography coining this “no-go theorem” into an application [Gisin et al. (2002)].

6.4.2.4 Elementary Quantum Algorithms

In the same scientific paper in which D. Deutsch introduced the notion of the universal quantum computer, he also presented the first quantum algorithm [Deutsch (1985)]. The problem that this algorithm addresses, later referred to as Deutsch’s problem, is a very simple one. Yet the Deutsch algorithm already exemplifies the advantages of a quantum computer through skillfully exploiting quantum parallelism. Like the Deutsch algorithm, all other elementary quantum algorithms amount to deciding which black box out of finitely many alternatives one has at hand. Such a black box is often also referred to as oracle. An input may be given to the oracle, one may read out or use the outcome in later steps of the quantum algorithm, and the objective is to find out the functioning of the black box. It is assumed that this oracle operation can be implemented with some sequence of quantum logic gates. The complexity of the quantum algorithm is then quantified in terms of the number of queries to the oracle [Eisert and Wolf (2006)].

With the help of the Deutsch algorithm, it is possible to decide whether a function has a certain property with a single call of the function, instead of two calls that are necessary classically.

Let \( f : \{0, 1\} \rightarrow \{0, 1\} \) be a function that has both a one-bit domain and range. This function can be either constant or balanced, which means that either \( f(0) \oplus f(1) = 0 \) or \( f(0) \oplus f(1) = 1 \) holds. The problem is to find out with the minimal number of function calls whether this function \( f \) is constant or balanced.

Classically, it is obvious that two function calls are required to decide...
which of the two allowed cases is realised, or, equivalently, what the value of \( f(0) \oplus f(1) \) is. A way to compute the function \( f \) on a quantum computer is to transform the state vector of two qubits according to

\[
| x, y \rangle \mapsto U_f | x, y \rangle = | x, f(x) \oplus y \rangle.
\]

(6.59)

In this manner, the evaluation can be realized unitarily. The above map is what is called a standard quantum oracle (as opposed to a minimal quantum oracle \[\text{Vedral (1996)}\], which would be of the form \(| x \rangle \mapsto | f(x) \rangle \)). The claim now is that using such an oracle, a single function call is sufficient for the evaluation of \( f(0) \oplus f(1) \). In order to show this, let us assume that we have prepared two qubits in the state with state vector \(| \Psi \rangle = (H \otimes H)|0, 1\rangle \), where \(H\) denotes the Hadamard gate. We now apply the unitary \(U_f\) once to this state, and finally apply another Hadamard gate to the first qubit. The resulting state vector hence reads as

\[
| \Psi' \rangle = (H \otimes 1)U_f(H \otimes H)|0, 1\rangle.
\]

(6.60)

A short calculation shows that \(| \Psi' \rangle\) can be evaluated to

\[
| \Psi' \rangle = \pm| f(0) \oplus f(1)\rangle H|1\rangle.
\]

(6.61)

The second qubit is in the state corresponding to the vector \(H|1\rangle\), which is of no relevance to our problem. The state of the first qubit, however, is quite remarkable: encoded is \(| f(0) \oplus f(1)\rangle\), and both alternatives are decidable with unit probability in a measurement in the computational basis, as the two state vectors are orthogonal. That is, with a single measurement of the state, and notably, with a single call of the function \(f\), of the first qubit we can decide whether \(f\) was constant or balanced.

For more details on quantum algorithms, see \[\text{Eisert and Wolf (2006)}\].

6.5 The Hardware for Quantum Computers

This section addresses modern electronic devices called Josephson junctions, which promise to be a basic building blocks of the future quantum computers. Apparently, they can exhibit chaotic behavior, both as single junctions (which have macroscopic dynamics analogous to those of the forced nonlinear oscillators), and as arrays (or ladders) of junctions, which can show high-dimensional chaos.

A Josephson junction is a type of electronic circuit capable of switching at very high speeds, i.e., frequency of typically \(10^{10} - 10^{11}\) Hz, when
operated at temperatures approaching absolute zero. It is an insulating barrier separating two superconducting materials and producing the Josephson effect. The terms are named eponymously after British physicist Brian David Josephson, who predicted the existence of the Josephson effect in 1962 [Josephson (1974)]. Josephson junction exploits the phenomenon of superconductivity, the ability of certain materials to conduct electric current with practically zero resistance. Josephson junctions have important applications in quantum–mechanical circuits. They have great technological promises as amplifiers, voltage standards, detectors, mixers, and fast switching devices for digital circuits. They are used in certain specialized instruments such as highly–sensitive microwave detectors, magnetometers, and QUIDs. Finally, Josephson junctions allow the realisation of qubits, the key elements of quantum computers.

Josephson junctions have been particularly useful for experimental studies of nonlinear dynamics as the equation governing a single junction dynamics is the same as that for a pendulum [Strogatz (1994)]. Their dynamics can be analyzed both in a simple overdamped limit and in the more complex underdamped one, either for single junctions and for arrays of large numbers of coupled junctions.

A Josephson junction is made up of two superconductors, separated by a weak coupling non–superconducting layer, so thin that electrons can cross through the insulating barrier. It can be conceptually represented as:

Superconductor 1 : $\psi_1 e^{i\phi_1}$
Weak Coupling $\leftrightarrow$
Superconductor 2 : $\psi_2 e^{i\phi_2}$

where the two superconducting regions are characterized by simple quantum–mechanical wave functions, $\psi_1 e^{i\phi_1}$ and $\psi_2 e^{i\phi_2}$, respectively. Normally, a much more complicated description would be necessary, as there are $\sim 10^{23}$ electrons to deal with, but in the superconducting ground state, these electrons form the so–called Cooper pairs that can be described by a single macroscopic wave function $\psi e^{i\phi}$. The flow of current between the superconductors in the absence of an applied voltage is called a Josephson current, and the movement of electrons across the barrier is known as Josephson tunnelling (see Figure 6.8). Two or more junctions joined by superconducting paths form what is called a Josephson interferometer.

One of the characteristics of a Josephson junction is that as the temperature is lowered, superconducting current flows through it even in the
The Josephson effect in particular results from two superconductors acting to preserve their long-range order across an insulating barrier. With a thin enough barrier, the phase of the electron wave–function in one superconductor maintains a fixed relationship with the phase of the wave–function in another superconductor. This linking up of phase is called phase coherence. It occurs throughout a single superconductor, and it occurs between the superconductors in a Josephson junction. The phase coherence, or long-range order, is the essence of the Josephson effect.

While researching superconductivity, B.D. Josephson studied the properties of a junction between two superconductors. Following up on earlier work by L. Esaki and I. Giaever, he demonstrated that in a situation when there is electron flow between two superconductors through an insulating layer (in the absence of an applied voltage), and a voltage is applied, the current stops flowing and oscillates at a high frequency. The Josephson effect is influenced by magnetic fields in the vicinity, a capacity that en-
ables the Josephson junction to be used in devices that measure extremely weak magnetic fields, such as superconducting quantum interference devices (SQUIDs). For their efforts, Josephson, Esaki, and Giaever shared the Nobel Prize for Physics in 1973.

The *Josephson–junction quantum computer* was demonstrated in April 1999 by Nakamura, Pashkin and Tsai of NEC Fundamental Research Laboratories in Tsukuba, Japan [Nakamura et al. (1999)]. In the same month, only about one week earlier, Ioffe, Geshkenbein, Feigel’man, Fauchère and Blatter, independently, described just such a computer in Nature [Ioffe et al. (1999)].

Nakamura, Pashkin and Tsai’s computer is built around a *Cooper pair box*, which is a small superconducting island electrode weakly coupled to a bulk superconductor. Weak coupling between the superconductors creates a Josephson junction between them. Like most other junctions, the Josephson junction is also a capacitor, which is charged by the current that flows through it. A gate voltage is applied between the two superconducting electrodes. If the Cooper box is sufficiently small, e.g., as small as a quantum dot, the charging current breaks into discrete transfer of individual Cooper pairs, so that ultimately it is possible to just transfer a single Cooper pair across the junction. The effectiveness of the Cooper pair transfer depends on the energy difference between the box and the bulk and a maximum is reached when a voltage is applied, which equalizes this energy difference. This leads to resonance and observable coherent quantum oscillations [Averin (1999)].

This contraption, like the Loss–Vincenzo quantum dot computer [Loss and DiVincenzo (1998)], has the advantage that it is controlled electrically. Unlike Loss–Vincenzo computer, this one actually exists in the laboratory. Nakamura, Pashkin and Tsai did not perform any computations with it though. At this stage it was enough of an art to observe the coherence for about 6 cycles of the Cooper pair oscillations, while the chip was cooled to about and carefully shielded from external electromagnetic radiation.

There are two general types of Josephson junctions: overdamped and underdamped. In overdamped junctions, the barrier is conducting (i.e., it is a normal metal or superconductor bridge). The effects of the junction’s internal electrical resistance will be large compared to its small capacitance. An overdamped junction will quickly reach a unique equilibrium state for any given set of conditions.

The barrier of an underdamped junction is an insulator. The effects of the junction’s internal resistance will be minimal. Underdamped junctions
do not have unique equilibrium states, but are hysteretic.

A Josephson junction can be transformed into the so-called Giaever tunnelling junction by the application of a small, well defined magnetic field. In such a situation, the new device is called a superconducting tunnelling junction (STJ) and is used as a very sensitive photon detector throughout a wide range of the spectrum, from infrared to hard X-ray. Each photon breaks up a number of Cooper pairs. This number depends on the ratio of the photon energy to approximately twice the value of the gap parameter of the material of the junction. The detector can be operated as a photon-counting spectrometer, with a spectral resolution limited by the statistical fluctuations in the number of released charges. The detector has to be cooled to extremely low temperature, typically below 1 kelvin, to distinguish the signals generated by the detector from the thermal noise. Small arrays of STJs have demonstrated their potential as spectro-photometers and could further be used in astronomy [ESA (2005)]. They are also used to perform energy dispersive X-ray spectroscopy and in principle they could be used as elements in infrared imaging devices as well [Enss (2005)].

6.5.1 Josephson Effect and Pendulum Analog

6.5.1.1 Josephson Effect

The basic equations governing the dynamics of the Josephson effect are (see, e.g., [Barone and Paterno (1982)]):

\[ U(t) = \frac{\hbar}{2e} \frac{\partial \phi}{\partial t}, \quad I(t) = I_c \sin \phi(t), \]

where \( U(t) \) and \( I(t) \) are the voltage and current across the Josephson junction, \( \phi(t) \) is the phase difference between the wave functions in the two superconductors comprising the junction, and \( I_c \) is a constant, called the critical current of the junction. The critical current is an important phenomenological parameter of the device that can be affected by temperature as well as by an applied magnetic field. The physical constant \( \hbar/2e \) is the magnetic flux quantum, the inverse of which is the Josephson constant.

The three main effects predicted by Josephson follow from these relations:

1. The DC Josephson effect. This refers to the phenomenon of a direct current crossing the insulator in the absence of any external electromagnetic field, owing to Josephson tunnelling. This DC Josephson current is proportional to the sine of the phase difference across the insulator, and
may take values between $-I_c$ and $I_c$.

2. The AC Josephson effect. With a fixed voltage $U_{DC}$ across the junctions, the phase will vary linearly with time and the current will be an AC current with amplitude $I_c$ and frequency $2e/hU_{DC}$. This means a Josephson junction can act as a perfect voltage-to-frequency converter.

3. The inverse AC Josephson effect. If the phase takes the form

$$\phi(t) = \phi_0 + n\omega t + a\sin(\omega t),$$

the voltage and current will be

$$U(t) = \frac{h}{2e}\omega[n + a\cos(\omega t)], \quad I(t) = I_c \sum_{m=-\infty}^{\infty} J_n(a) \sin[\phi_0 + (n + m)\omega t].$$

The DC components will then be

$$U_{DC} = n\frac{h}{2e}\omega, \quad I(t) = I_c J_{-n}(a) \sin \phi_0.$$

Hence, for distinct DC voltages, the junction may carry a DC current and the junction acts like a perfect frequency-to-voltage converter.

6.5.1.2 Pendulum Analog

To show a driven pendulum analog of a microscopic description of a single Josephson junction, we start with:

1. The Josephson current–phase relation

$$I = I_c \sin \phi,$$

where $I_c$ is the critical current, $I$ is the bias current, and $\phi = \phi_2 - \phi_1$ is the constant phase difference between the phases of the two superconductors that are weakly coupled; and

2. The Josephson voltage–phase relation

$$V = \frac{h}{2e} \dot{\phi},$$

where $V = V(t)$ is the instantaneous voltage across the junction, $h$ is the Planck constant (divided by $2\pi$), and $e$ is the charge on the electron.

Now, if we apply Kirchhoff’s voltage and current laws for the parallel RC–circuit with resistance $R$ and capacitance $C$, we come to the first–order ODE

$$C\ddot{V} + \frac{V}{R} + I_c \sin \phi = I,$$
which can be recast solely in terms of the phase difference $\phi$ as the second-order pendulum–like ODE,

\[
\text{Josephson junction : } \frac{\hbar C}{2e} \dddot{\phi} + \frac{\hbar}{2eR} \ddot{\phi} + I_c \sin \phi = I, \quad (6.62)
\]

\[
\text{Pendulum : } ml^2 \ddot{\theta} + b\dot{\theta} + mgl \sin \theta = \tau.
\]

This mechanical analog has often proved useful in visualizing the dynamics of Josephson Junctions [Strogatz (1994)]. If we divide (6.62) by $I_c$ and define a dimensionless time

\[
\tau = \frac{2eI_cR}{\hbar} t,
\]

we get the dimensionless oscillator equation for Josephson junction,

\[
\beta \phi'' + \phi' + \sin \phi = \frac{I}{I_c}, \quad (6.63)
\]

where $\phi' = d\phi/d\tau$. The dimensionless group $\beta$, defined by

\[
\beta = \frac{2eI_cR^2C}{\hbar},
\]

is called the McCumber parameter and represents a dimensionless capacitance.

In a simple overdamped limit $\beta \ll 1$ with resistive loading, the ‘inertial term’ $\beta \phi''$ may be neglected (as if oscillating in a highly-viscous medium), and so (6.63) reduces to a non-uniform oscillator

\[
\phi' = \frac{I}{I_c} - \sin \phi, \quad (6.64)
\]

with solutions approaching a stable fixed-point for $I < I_c$, and periodically varying for $I < I_c$. To find the current–voltage curve in the overdamped limit, we take the average voltage $\langle V \rangle$ as a function of the constant applied current $I$, assuming that all transients have decayed and the system has reached steady-state, and get

\[
\langle V \rangle = I_c R \langle \phi' \rangle.
\]

An overdamped array of $N$ Josephson Junctions [6.64], parallel with a resistive load $R$, can be described by the system of first-order dimensionless
ODEs [Strogatz (1994)]

\[
\phi_k' = \Omega + a \sin \phi_k + \frac{1}{N} \sum_{j=1}^{N} \sin \phi_j, \quad k = 1, \ldots, N,
\]

where

\[
\Omega = I_b R_0 / I_c r, \quad a = -(R_0 + r)/r, \quad R_0 = R/N,
\]

\[
I_b = I_c \sin \phi_k + \frac{\hbar}{2eR} \dot{\phi}_k + \frac{\hbar}{2eR} \sum_{j=1}^{N} \dot{\phi}_j.
\]

**6.5.2 Dissipative Josephson Junction**

The past decade has seen a considerable interest and remarkable activity in an area which presently is often referred to as macroscopic quantum mechanics. Specifically, one has been interested in quantum phenomena of macroscopic objects [Leggett (1986)].

In particular, macroscopic quantum tunnelling [Caldeira and Leggett (1981)] (quantum decay of a meta–stable state), and quantum coherence [Leggett et al. (1987)] have been studied. Soon, it became clear that dissipation has a profound influence on these quantum phenomena. Phenomenologically, dissipation is the consequence of an interaction of the object with an environment which can be thought of as consisting of infinitely many degrees of freedom. Specifically, the environmental degrees of freedom may be chosen to be harmonic oscillators such that we may consider the dissipation as a process where excitations, that are phonons, are emitted and absorbed. This, Caldeira–Leggett model has been used in [Caldeira and Leggett (1981)] where the influence of dissipation on tunnelling has been explored.

As far as quantum coherence is concerned, the most simple system is an object with two different quantum states: it is thought to represent the limiting case of an object in a double-well potential where only the lowest energy states in each of the two wells is relevant and where the tunnelling through the separating barrier allows for transitions that probe the coherence. Since a 2–state system is equivalent to a spin–one–half problem, this standard system is often referred to by this name. In particular, with the standard coupling to a dissipative environment made of harmonic oscillators, it is called the spin–boson problem which has been studied repeatedly in the past [Leggett et al. (1987)].
Level quantization and resonant tunnelling have been observed recently [Van der Vaart et al. (1995)] in a double-well quantum-dot system. However, the influence of dissipation was not considered in this experiment. On the other hand, it seems that Josephson junctions are also suitable systems for obtaining experimental evidence pertaining to macroscopic quantum effects. In this context, evidence for level quantization and for quantum decay have been obtained [Martinis et al. (1985)].

Recall that a Josephson junction may be characterized by a current–phase relation

\[ I(\phi) = I_J \sin \phi, \]  \hspace{1cm} (6.65)

where the phase \( \phi \) is related to the voltage difference \( U \) by

\[ \hbar \dot{\phi} = 2eU. \]  \hspace{1cm} (6.66)

Therefore, the phase of a Josephson junction shunted by a capacitance \( C \) and biased by an external current \( I_x \) obeys a classical type of equation of motion

\[ M \ddot{\phi} = -\frac{\partial V(\phi)}{\partial \phi}, \quad \text{with the mass} \]  \hspace{1cm} (6.67)

\[ M = \left( \frac{\hbar}{2e} \right)^2 C, \quad \text{and the potential energy} \]  \hspace{1cm} (6.68)

\[ V(\phi) = -\frac{\hbar}{2e} \left[ I_J \cos \phi + I_x \phi \right]. \]  \hspace{1cm} (6.69)

A widely discussed model of a dissipative object is the one where the Josephson junction is also shunted by an Ohmic resistor \( R \). In this case, the classical equation of motion (6.67) has to be replaced by

\[ M \ddot{\phi} = -\frac{\partial V(\phi)}{\partial \phi} - \eta \dot{\phi}, \quad \eta = \left( \frac{\hbar}{2e} \right)^2 \frac{1}{R}. \]  \hspace{1cm} (6.70)

The model of a dissipative environment according to the above specification has been discussed by [Caldeira and Leggett (1981)].

The potential energy \( \tilde{V}(\phi) \) of (6.69) displays wells at \( \phi \approx 2n\pi \) with depth shifted by an amount \( \Delta \approx (2\pi\hbar/2e)I_x \). If the wells are sufficiently deep, one needs to concentrate only on transitions between pairs of adjacent wells. Thus, one arrives at the double well problem mentioned above.
The analysis in this paper goes beyond the limiting situation where only the lowest level in each of the two wells is of importance. Roughly, this is realized when the level separation $\hbar (2E_J/M)^{1/2} \approx (2eI_J/C)^{1/2}$ is smaller than or comparable with $\Delta$. In particular, we will concentrate on resonance phenomena which are expected to show up whenever two levels in the adjacent wells happen to cross when the bias current $I_x$, that is $\Delta$, is varied.

For such values of the bias current, there appear sharp asymmetric peaks in the current-voltage characteristic of the Josephson junction. This phenomenon has been studied by Larkin et al. (1988) within the standard model in the one–phonon approximation. For bias currents that correspond to crossings of the next and next nearest levels (e.g., ground state in the left well and the first or second excited state at the right side), it is possible to neglect processes in the reverse direction provided that the temperature is sufficiently low. Thus, the restriction to a double well system receives additional support.

The transfer of the object from the left to the right potential well is accompanied by the emission of an infinite number of phonons. Therefore, in Ovchinnikov and Schmid (1994) the fact is taken into account that in the resonance region, the contribution of phonons of small energy is important as well as the contribution of resonance phonons with energy equal to the distance between levels in the wells.

6.5.2.1 Junction Hamiltonian and its Eigenstates

The model of Müllers and Schmid (1995) consists of a particle, called ‘object’ (coordinate $R_1$), which is coupled (in the sense of Caldeira and Leggett (1981)) to a ‘bath’ of harmonic oscillators (coordinates $R_j$). We shall use the conventions $j \in \{2, \ldots, N\}$ for the bath oscillators and $k \in \{1, \ldots, N\}$ for the indices of all coordinates in the model. The double–well potential is approximated by two parabolas about the minima of the two wells.

The phase $\phi$ of the Josephson contact then corresponds to the object coordinate $R_1$ of the model, and the voltage $U$ is related to the tunnelling rate $J$ by $2eU = \dot{\phi} = 2\pi J$. As it has already been remarked, the current $I_x$ is proportional to the bias $\Delta$ of the two wells. Thus, calculating the transition rate for different values of the bias $\Delta$ is equivalent to the determination of the I–V characteristics.

Specifically, following Müllers and Schmid (1995), we want to write
the Hamiltonian of the model in the form
\[
\hat{H} = \frac{1}{2m} \sum_k \hat{p}_k^2 + \hat{v}(\hat{R}_1) + \frac{m}{2} \sum_j \omega_j^2 (\hat{R}_j - \hat{R}_1)^2,
\]
\[
\hat{v}(\hat{R}_1) \approx \frac{m}{2} \sum_{\pm} \Omega^2 (\hat{R}_1 \pm a)^2 \pm \Delta^2.
\tag{6.71}
\]

The states for the two situations ‘object in the left well’ and ‘object in the right well’ will be denoted by \(|\Lambda_L, L\rangle\) and \(|\Lambda_R, R\rangle\), respectively. If one projects onto the eigenstates \(|n\rangle\) of the 1D harmonic oscillator and takes into account the shift of the wells, one arrives at the following decomposition
\[
\langle n_L, \{R_j\}|\Lambda_L, L\rangle = \int dR_1 \phi_{n_L}(R_1 + a) \phi^L_{\Lambda_L}(\{R_k\}),
\tag{6.72}
\]
\[
\langle n_R, \{R_j\}|\Lambda_R, R\rangle = \int dR_1 \phi_{n_R}(R_1 - a) \phi^R_{\Lambda_R}(\{R_k\}).
\]

The situations ‘object on the left’ and ‘object on the right’ differ only by the shift and the bias of the wells. Therefore, one can find a unified representation by noting that \(\phi^L_{\Lambda_L}(\{R_k\}) = \Phi^L_{\Lambda}(\{R_k + a\})\) and \(\phi^R_{\Lambda_R}(\{R_k\}) = \Phi^R_{\Lambda}(\{R_k - a\})\). The eigenstates \(\Phi_{\Lambda}\) are defined by the relations
\[
\Phi_{\Lambda}(\{R_k\}) = \langle \{R_k\}|\Lambda\rangle, \quad \hat{H}_0|\Lambda\rangle = E_{\Lambda}|\Lambda\rangle,
\]
\[
\hat{H}_0 = \frac{1}{2m} \sum_k \hat{p}_k^2 + \frac{m}{2} \sum_j \omega_j^2 (\hat{R}_j - \hat{R}_1)^2 + \frac{m}{2} \Omega^2 \hat{R}_1^2.
\]

Thus, it follows from (6.72) that
\[
\langle n_L, \{R_j\}|\Lambda_L, L\rangle = \langle n_L, \{R_j\}|\exp(ia \sum_j \hat{p}_j)|\Lambda_L\rangle, \tag{6.73}
\]
\[
\langle n_R, \{R_j\}|\Lambda_R, R\rangle = \langle n_R, \{R_j\}|\exp(-ia \sum_j \hat{p}_j)|\Lambda_R\rangle,
\]

where we have used the shift property of the momentum operator \(\hat{p}\).

The coupling of the two wells is taken into account by means of a tunnelling Hamiltonian \(\hat{H}_T\) which we represent in the form
\[
\langle \Lambda_L, L|\hat{H}_T|\Lambda_R, R\rangle = \int d\{R_j\} \sum_{n_L,n_R} T_{n_L,n_R} \langle \Lambda_L, L|n_L, \{R_j\}\rangle \langle n_R, \{R_j\}|\Lambda_R, R\rangle.
\]

Using again the momentum operator, one can write
\[
|x\rangle \langle x'| = e^{i\hat{p}(x'-x)} |x\rangle \langle x'| = e^{i\hat{p}(x'-x)} \delta(x' - \hat{x}).
\]
From this, we conclude that

\[ \langle \Lambda_L, L | \hat{H}_T | \Lambda_R, R \rangle = \sum_{nL,nR} T_{nL,nR} \int dR_1 dR'_1 \frac{d\mathcal{Q}}{2\pi} \phi^*_n(R'_1) \phi_n(R_1) \times \langle \Lambda_L, L | e^{i\hat{A}(R'_1-R_1)} e^{i\mathcal{Q}(R'_1-R_1)} | \Lambda_R, R \rangle. \]

6.5.2.2 Transition Rate

The net transition rate from the left well to the right one is then in second order perturbation theory given by [Müllers and Schmid (1995)]

\[ J = 2\pi Z_0^{-1} \sum_{\Lambda_L,\Lambda_R} |\langle \Lambda_L, L | \hat{H}_T | \Lambda_R, R \rangle|^2 \delta(E_{\Lambda_L} - E_{\Lambda_R} + \Delta) \times [e^{-\beta E_{\Lambda_L}} - e^{\beta E_{\Lambda_R}}], \]

where \( Z_0 = \text{Tr} \exp(-\beta \hat{H}_0) \). The \( \delta \)-function may be written in Fourier representation, and the fact that the \( E_{\Lambda} \) are eigen–energies of \( \hat{H}_0 \) serves us to incorporate the energy conservation into Heisenberg time–dependent operators \( \hat{A}(t) = \exp(i\hat{H}_0 t) \hat{A} \exp(-i\hat{H}_0 t) \), i.e.,

\[ \langle \Lambda_L, L | \hat{H}_T | \Lambda_R, R \rangle \delta(E_{\Lambda_L} - E_{\Lambda_R} + \Delta) = \int dt e^{i\Delta t} \langle \Lambda_L | e^{-i\Delta \sum_k \hat{A}_k(t)} \hat{H}_T(t) e^{-i\Delta \sum_k \hat{A}_k(t)} | \Lambda_R \rangle. \]

Then, collecting our results from above we arrive at the expression

\[ J = Z_0^{-1} (1 - e^{-\beta \Delta}) \int dt e^{i\Delta t} \sum_{nL,nR} T_{nL,nR} T^\ast_{nL,nR} \]

\[ \times \int \frac{d\mathcal{Q}d\mathcal{Q}'}{(2\pi)^2} \int dR_1 dR'_1 d\mathcal{R}_1 d\mathcal{R}'_1 \phi^*_n(R'_1) \phi_n(R_1) \phi^*_n(R'_1) \phi_n(R_1) \]

\[ \times \text{Tr} \left\{ e^{-\beta \hat{H}_0} e^{-i\Delta \sum_k \hat{A}_k(t)} e^{i\hat{A}_k(t)(R'_1-R_1+2a)} e^{-i\mathcal{Q}(R'_1-R_1)} \right\}. \]

Let us now use the relation

\[ e^{-i(H_0+\mathcal{W})t} = e^{-i\hat{H}_0 t} \hat{T} e^{-i \int_0^t dt' \hat{W}(t')} \]

which holds for \( t > 0 \) when \( \hat{T} \) is the time–ordering operator and for \( t < 0 \) when the anti time–ordering is used. If we define \( \langle \hat{A} \rangle = \text{Tr} \exp(-\beta \hat{H}_0) \hat{A} / Z_0 \), we can write the following result for the transition
rate:
\[
J = (1 - e^{-\beta \Delta}) \int dt \sum_{n_L, n_R} T_{n_L, n_R} T_{n_L, n_R}^{*} \exp \left[ i \int dt \left( \Delta - 2m\Omega^2 a^2 t \right) + i \int dt \left( \Sigma_{-} - \Sigma_{+} \right) + i \int dt \left( \Sigma_{0} \right) \right] \sum_{n_L, n_R} T_{n_L, n_R} T_{n_L, n_R}^{*}.
\]

We are now in the position to make use of the fact that the Hamiltonian is quadratic in all coordinates so that we can evaluate exactly
\[
\langle \hat{T} e^{i \int dt \eta(t)} \hat{R}(t) \rangle = e^{-i \int dt \left( \int dQ dQ' \frac{1}{(2\pi)^2} \int dR dR' \phi_{n_L}(R) \phi_{n_R}^{*}(R') \phi_{n_L}(R) \phi_{n_R}^{*}(R') \exp \left[ -i Q \hat{R}(t) + i \hat{p}(t) + 2im\Omega^2 a \int_{0}^{t} dt' \hat{R}(t') - i Q \hat{R}(0) + i \hat{p}(0) \right] \int dt' \hat{R}(t') \rangle \quad (6.74)
\]

By comparison with the last two lines in eq. (6.74), the function \( \eta(t') \) is given by
\[
\eta(t') = -Q \delta(t' - t) - Q \delta(t') + 2m\Omega^2 a \left( \Theta(t') - \Theta(t' - t) \right)
\]
\[
+ m(R - R') \delta'(t' - t) + m(\overline{R} - \overline{R}') \delta'(t') .
\]

\( \Theta(t) \) is meant to represent the step function. The derivatives of the \( \delta \)-function arise from a partial integration of terms containing \( \hat{p}(t) = m \frac{d}{dt} \hat{x}(t) \). Note, that these act only on the coordinates but not on the step functions which arise due to the time ordering.

Moreover, the degrees of freedom of the bath can be integrated out in the usual way \cite{Caldeira and Leggett 1981} leading to a dissipative influence on the object. One is then lead to the following form of the Fourier transform of \( D(t, t') \equiv D(t - t') \):
\[
D(\omega) = \frac{D_R(\omega)}{1 - \exp \left( -h\omega/k_B T \right)} + \frac{D_R(-\omega)}{1 - \exp \left( h\omega/k_B T \right)} , \quad (6.76)
\]
\[
(D_R^{-1})(\omega) = m[(\omega + i0)^2 - \Omega^2] + i\omega ,
\]
where we will use a spectral density \( J(\omega) = \eta \omega, 0 \leq \omega \leq \omega_c \) for the bath oscillators.

From \( (6.74) \) and \( (6.75) \) one can conclude that the integrations with respect to \( \xi, \xi', R, R', \overline{R}, \overline{R}' \) can be done exactly as only Gaussian integrals...
are involved (note, that the eigenstates of the harmonic oscillator are Gaussian functions and derivatives of these, respectively). Therefore, for given $n_L, n_R, \pi_L, \pi_R$, one has to perform a 6D Gaussian integral [Müller and Schmid (1995)].

### 6.5.3 Josephson Junction Ladder (J JL)

2D arrays of Josephson junctions have attracted much recent theoretical and experimental attention. Interesting physics arises as a result of competing vortex-vortex and vortex-lattice interactions. It is also considered to be a convenient experimental realization of the so-called frustrated XY models. Here, mainly following [Denniston and Tang (1995)], we discuss the simplest such system, namely the Josephson junction ladder (J JL, see Figure 6.9) [Kardar (1984); Granato (1990)].

![Figure 6.9 Josephson junction Ladder (J JL).](image)

To construct the system, superconducting elements are placed at the ladder sites. Below the bulk superconducting-normal transition temperature, the state of each element is described by its charge and the phase of the superconducting wave $\psi$-function [Anderson (1964)]. In this section we neglect charging effects, which corresponds to the condition that $4e^2/C \ll J$, with $C$ being the capacitance of the element and $J$ the Josephson coupling. Let $\theta_j$ ($\theta'_j$) denote the phase on the upper (lower) branch of the ladder at the jth rung. The Hamiltonian for the array [Tinkham (1975)] can be written in terms the gauge invariant phase differences [Denniston]...
\[\gamma_j = \theta_j - \theta_{j-1} - (2\pi/\phi_0) \int_{j-1}^j A_x dx, \quad \gamma_j' = \theta_j' - \theta_{j-1}' - (2\pi/\phi_0) \int_{j-1}^{j'} A_x dx,\]

and \[\alpha_j = \theta_j' - \theta_j - (2\pi/\phi_0) \int_j^{j'} A_y dx\]

as

\[H = -\sum_j (J_x \cos \gamma_j + J_x \cos \gamma_j' + J_y \cos \alpha_j),\]  

(6.77)

where \(A_x\) and \(A_y\) are the components of the magnetic vector potential along and transverse to the ladder, respectively, and \(\phi_0\) the flux quantum. The sum of the phase differences around a plaquette is constrained by

\[\gamma_j - \gamma_j' + \alpha_j - \alpha_{j-1} = 2\pi(f - n_j),\]

where \(n_j = 0, \pm 1, \pm 2, \ldots\) is the vortex occupancy number and \(f = \phi/\phi_0\) with \(\phi\) being the magnetic flux through a plaquette. With this constraint, it is convenient to write (6.77) in the form

\[H = -J \sum_j \left\{ 2 \cos \eta_j \cos [\frac{(\alpha_{j-1} - \alpha_j)}{2} + \pi (f - n_j)] + J_t \cos \alpha_j \right\},\]  

(6.78)

where \(\eta_j = (\gamma_j + \gamma_j')/2, J = J_x\) and \(J_t = J_y/J_x\).

The Hamiltonian is symmetric under \(f \rightarrow f + 1\) with \(n_j \rightarrow n_j + 1\), and \(f \rightarrow -f\) with \(n_j \rightarrow -n_j\), thus it is sufficient to study only the region \(0 \leq f \leq 0.5\). Since in one dimension ordered phases occur only at zero temperature, the main interest is in the ground states of the ladder and the low temperature excitations. Note that in (6.78) \(\eta_j\) decouples from \(\alpha_j\) and \(n_j\), so that all the ground states have \(\eta_j = 0\) to minimize \(H\). The ground states will be among the solutions to the current conservation equations:

\[\partial_{\alpha_j} H = 0, \text{ i.e., } \quad \text{[Denniston and Tang (1995)]}\]

\[J_t \sin \alpha_j = \sin [(\alpha_{j-1} - \alpha_j)/2 + \pi (f - n_j)] - \sin [(\alpha_j - \alpha_{j+1})/2 + \pi (f - n_{j+1})].\]  

(6.79)

For any given \(f\) there are a host of solutions to (6.79). The solution that minimizes the energy must be selected to get the ground state.

If one expands the inter–plaquette cosine coupling term in (6.78) about it’s maximum, the discrete sine–Gordon model is obtained. A vortex \((n_j = 1)\) in the J JL corresponds to a kink in the sine–Gordon model. This analogy was used by [Kardar (1984)] as an argument that this
Quantum Leap

system should show similar behavior to the discrete sine–Gordon model which has been studied by several authors [Aubry and André (1980), Coppersmith and Fisher (1983), Pokrovsky et al. (1986)]. This analogy is only valid for $J_t$ very small so that the inter–plaquette term dominates the behavior of the system making the expansion about its maximum a reasonable assumption. However, much of the interesting behavior of the discrete sine–Gordon model occurs in regions of large $J_t$ ($J_t \sim 1$). Furthermore, much of the work by Aubry [Aubry and André (1980)] on the sine–Gordon model relies on the convexity of the coupling potential which we do not have in the JJL.

Following Denniston and Tang (1995), here we formulate the problem in terms of a transfer matrix obtained from the full partition function of the ladder. The eigenvalues and eigenfunctions of the transfer matrix are found numerically to determine the phases of the ladder as functions of $f$ and $J_t$. We study the properties of various ground states and the low temperature excitations. As $J_t$ is varied, all incommensurate ground states undergo a superconducting–normal transition at certain $J_t$ which depends on $f$. One such transition will be analyzed. Finally we discuss the critical current.

The partition function for the ladder, with periodic boundary conditions and $K = J/k_B T$, is

$$Z = \prod_i \int_{-\pi}^{\pi} d\alpha_i d\eta_i \exp\{K(2 \cos \eta_i \cos[(\alpha_{i-1} - \alpha_i)/2 + \pi(f - n_i)] + J_t \cos \alpha_i)\}.$$  

The $\eta_i$ can be integrated out resulting in a simple transfer matrix formalism for the partition function involving only the transverse phase differences:

$$Z = \prod_i \int_{-\pi}^{\pi} d\alpha_i P(\alpha_{i-1}, \alpha_i) = Tr \hat{P}^N.$$  

The transfer matrix elements $P(\alpha, \alpha')$ are

$$P(\alpha, \alpha') = 4\pi \exp[KJ_t(\cos \alpha + \cos \alpha')/2] I_0(2K \cos[(\alpha - \alpha')/2 + \pi f]),$$  \hspace{1cm} (6.80)

where $I_0$ is the zeroth order modified Bessel function. Note that the elements of $\hat{P}$ are real and positive, so that its largest eigenvalue $\lambda_0$ is real, positive and nondegenerate. However, since $\hat{P}$ is not symmetric (except for $f = 0$ and $f = 1/2$) other eigenvalues can form complex conjugate pairs.
As we will see from the correlation function, these complex eigenvalues determine the spatial periodicity of the ground states.

The two point correlation function of $\alpha_j$'s is \cite{Denniston and Tang (1995)}

$$
\langle e^{i(\alpha_0 - \alpha_l)} \rangle = \lim_{N \to \infty} \frac{\prod_i^N \int_{-\pi}^{\pi} d\alpha \rho(\alpha_{i-1}, \alpha_i) e^{i(\alpha_0 - \alpha_l)}}{Z} = \sum_n c_n (\frac{\lambda_n}{\lambda_0})^l,
$$

(6.81)

where we have made use of the completeness of the left and right eigenfunctions. (Note that since $\hat{P}$ is not symmetric both right $\psi^R_n$ and left $\psi^L_n$ eigenfunctions are need for the evaluation of correlation functions.) The $\lambda_n$ in (6.81) are the eigenvalues ($|\lambda_n| \geq |\lambda_{n+1}|$ and $n = 0, 1, 2, ...$), and the constants

$$
c_n = \int_{-\pi}^{\pi} d\alpha' \psi^L_0(\alpha') e^{i\alpha'} \psi^R_n(\alpha') \int_{-\pi}^{\pi} d\alpha \psi^L_n(\alpha) e^{-i\alpha} \psi^R_0(\alpha).
$$

In the case where $\lambda_1$ is real and $|\lambda_1| > |\lambda_2|$, (6.81) simplifies for large $l$ to

$$
\langle e^{i(\alpha_0 - \alpha_l)} \rangle = c_0 + c_1 \left( \frac{\lambda_1}{\lambda_0} \right)^l, \quad |\lambda_1| > |\lambda_2|.
$$

In the case where $\lambda_1 = \lambda_2 = |\alpha|^2 \Xi$, (6.81) for large $l$ is

$$
\langle e^{i(\alpha_0 - \alpha_l)} \rangle = c_0 + (c_1 e^{i2\pi \Xi l} + c_2 e^{-i2\pi \Xi l}) \left| \frac{\lambda_1}{\lambda_0} \right|^l, \quad \lambda_1 = \lambda_2.
$$

There is no phase coherence between upper and lower branches of the ladder and hence no superconductivity in the transverse direction. In this case, we say that the $\alpha$'s are unpinned. If there exist finite intervals of $\alpha$ on which $\rho(\alpha) = 0$, there will be phase coherence between the upper and lower branches and we say that the $\alpha$'s are pinned. In term of the transfer matrix, the phase density is the product of the left and right eigenfunctions of $\lambda_0$ \cite{Guyer and Miller (1979)},

$$
\rho(\alpha) = \psi^L_0(\alpha) \psi^R_0(\alpha).
$$

\footnote{While the correlation length is given by $\xi = [\ln|\lambda_0|/\lambda_1|]^{-1}$ the quantity $\Xi = Arg(\lambda_1)/2\pi$ determines the spatial periodicity of the state. By numerical calculation of $\lambda_n$, it is found that for $f$ smaller than a critical value $f_{c1}$ which depends on $J_t$, both $\lambda_1$ and $\lambda_2$ are real. These two eigenvalues become degenerate at $f_{c1}$, and then bifurcate into a complex conjugate pair \cite{Denniston and Tang (1995)}.}
We first discuss the case where $f < f_{c1}$. These are the Meissner–states in the sense that there are no vortices ($n_i = 0$) in the ladder. The Meissner–states are characterized by the absence of vortices. The ground state is simply $\alpha_i = 0$, $\gamma_j = \pi f$ and $\gamma_j' = -\pi f$, so that there is a global screening current $\pm J_x \sin \pi f$ in the upper and lower branches of the ladder [Kardar (1984)]. The phase density $\rho(\alpha) = \delta(\alpha)$. The properties of the Meissner state can be studied by expanding (6.78) around $\alpha_i = 0$,

$$H_M = (J/4) \sum_j [\cos(\pi f)(\alpha_{j-1} - \alpha_j)^2 + 2Jt\alpha_i^2].$$

The current conservation (6.79) becomes

$$\alpha_{j+1} = 2(1 + Jt/\cos \pi f) \alpha_j - \alpha_{j-1}. \quad (6.82)$$

Besides the ground state $\alpha_i = 0$, there are other two linearly independent solutions $\alpha_j = e^{\pm j/\xi_M}$ of (6.82) which describe collective fluctuations about the ground state, where

$$\frac{1}{\xi_M} = \ln \left[ 1 + \frac{Jt}{\cos \pi f} + \sqrt{\frac{2Jt}{\cos \pi f} + \left( \frac{Jt}{\cos \pi f} \right)^2} \right]. \quad (6.83)$$

$\xi_M$ is the low temperature correlation length for the Meissner state. As $f$ increases, the Meissner state becomes unstable to the formation of vortices. A vortex is constructed by patching the two solutions of (6.82) together using a matching condition. The energy $\epsilon_v$ of a single vortex is found to be

$$\epsilon_v \approx [2 + (\pi^2/8) \tanh(1/2\xi_M)] \cos \pi f - (\pi + 1) \sin \pi f + 2Jt, \quad (6.84)$$

for $J_t$ close to one. The zero of $\epsilon_v$ determines $f_{c1}$ which is in good agreement with the numerical result from the transfer matrix. For $f > f_{c1}$, $\epsilon_v$ is negative and vortices are spontaneously created. When vortices are far apart their interaction is caused only by the exponentially small overlap. The corresponding repulsion energy is of the order $J \exp(-l/\xi_M)$, where $l$ is the distance between vortices. This leads to a free energy per plaquette of $F = \epsilon_v/l + J \exp(-l/\xi_M)/l$ [Pokrovsky et al. (1986)]. Minimizing this free energy as a function of $l$ gives the vortex density for $f > f_{c1}$: $\langle n_j \rangle = l^{-1} = ([\xi_M \ln |f_{c1} - f|]^{-1}$ where a linear approximation is used for $f$ close to $f_{c1}$.

---

29 Here, $\xi_M < 1$ for $J_t \sim 1$ making a continuum approximation invalid.
We now discuss the commensurate vortex states, taking the one with \( \Xi = 1/2 \) as an example. This state has many similarities to the Meissner state but some important differences. The ground state is

\[
\alpha_0 = \arctan \left( \frac{2}{J_t} \sin(\pi f) \right), \quad \alpha_1 = -\alpha_0, \quad \alpha_{i \pm 2} = \alpha_i;
\]

\[n_0 = 0, \quad n_1 = 1, \quad n_{i \pm 2} = n_i, \quad (6.85)\]

so that there is a global screening current in the upper and lower branches of the ladder of \( \pm 2\pi J (f - 1/2) \sqrt{4 + J_t^2} \). The existence of the global screening, which is absent in an infinite 2D array, is the key reason for the existence of the steps at \( \Xi = p/q \). It is easy to see that the symmetry of this vortex state is that of the (antiferromagnetic) Ising model. The ground state is two-fold degenerate. The low temperature excitations are domain boundaries between the two degenerate ground states. The energy of the domain boundary \( \epsilon_b \) can be estimated using similar methods to those used to derive (6.84) for the Meissner state. We found that \( \epsilon_b = \epsilon_0^b - (\pi^2 / \sqrt{4 + J_t^2}) \vert f - 1/2 \vert \), where \( \epsilon_0^b \) depends only on

\[
J_t = c \arctan^2 (2/J_t) J_t^2 \coth (1/\xi_b) / \sqrt{4 + J_t^2},
\]

with \( c \) being a constant of order one and

\[\xi_b^{-1} = \ln(1 + J_t^2 / 2 + J_t \sqrt{1 + J_t^2 / 4}).\]

Thus the correlation length diverges with temperature as \( \xi \sim \exp(2J\epsilon_b/k_B T) \). The transition from the \( \Xi = 1/2 \) state to nearby vortex states happens when \( f \) is such that \( \epsilon_b = 0 \); it is similar to the transition from the Meissner state to its nearby vortex states. All other steps \( \Xi = p/q \) can be analyzed similarly. For comparison, we have evaluated \( \xi \) for various values of \( f \) and \( T \) from the transfer matrix and found that \( \xi \) fits \( \xi \sim \exp(2J\epsilon_b/k_B T) \) (typically over several decades) at low temperature.

We now discuss the superconducting–normal transition in the transverse direction. For \( J_t = 0 \), the ground state has \( \gamma_i = \gamma'_i = 0 \) and

\[\alpha_j = 2\pi f j + \alpha_0 - 2\pi \sum_{i=j}^{i=j} n_i. \quad (6.86)\]

The average vortex density \( \langle n_j \rangle \) is \( f \); there is no screening of the magnetic field. \( \alpha_0 \) in (6.86) is arbitrary; the \( \alpha \)'s are unpinned for all \( f \). The system is simply two un–coupled 1D XY chains, so that the correlation length \( \xi = 1/k_B T \). The system is superconducting at zero temperature along
the ladder, but not in the transverse direction. As \( J_t \) rises above zero we observe a distinct difference between the system at rational and irrational values of \( f \). For \( f \) rational, the \( \alpha \)'s become pinned for \( J_t > 0 \) (\( \rho(\alpha) \) is a finite sum of delta functions) and the ladder is superconducting in both the longitudinal and transverse directions at zero temperature. The behavior for irrational \( f \) is illustrated in the following for the state with \( \Xi = a_g \approx 0.381966 \cdots \) is one minus the inverse of the golden mean.

Finally, we consider critical currents along the ladder. One can get an estimate for the critical current by performing a perturbation expansion around the ground state (i.e., \( \{\alpha_j\} \) remain fixed) and imposing the current constraint of \( \sin \gamma_j + \sin \gamma_j' = I \). Let \( \delta \gamma_j, \delta \gamma_j' \) and \( \delta \alpha_j \) be the change of \( \gamma_j, \gamma_j' \) and \( \alpha_j \) in the current carrying state. One finds that stability of the ground state requires that \( \delta \alpha_j = 0 \), and consequently \( \delta \gamma_j = \delta \gamma_j' = I/2 \cos \gamma_j \). The critical current can be estimated by the requirement that the \( \gamma_j \) do not pass through \( \pi/2 \), which gives \( I_c = 2(\pi/2 - \gamma_{\text{max}}) \cos \gamma_{\text{max}} \), where \( \gamma_{\text{max}} = \max_j (\gamma_j) \). In all ground states we examined, commensurate and incommensurate, we found that \( \gamma_{\text{max}} < \pi/2 \), implying a finite critical current for all \( f \). See [Denniston and Tang (1995)] for more details.

6.5.3.1 Underdamped JJL

Recall that the discrete sine–Gordon equation has been used by several groups to model so-called hybrid Josephson ladder arrays [Ustinov et al. (1993); Watanabe et al. (1995)]. Such an array consists of a ladder of parallel Josephson junctions which are inductively coupled together (e.g., by superconducting wires). The sine-Gordon equation then describes the phase differences across the junctions. In an applied magnetic field, this equation predicts remarkably complex behavior, including flux flow resistance below a certain critical current, and a field–independent resistance above that current arising from so–called whirling modes [Watanabe et al. (1995)]. In the flux flow regime, the fluxons in this ladder propagate as localized solitons, and the IV characteristics exhibit voltage plateaus arising from the locking of solitons to linear spin–wave modes. At sufficiently large values of the anisotropy parameter \( \eta J \) defined later, the solitons may propagate ‘ballistically’ on the plateaus, i.e., may travel a considerable distance even after the driving current is turned off.

Here, mainly following [Ryu et al. (1996)], we show that this behavior is all found in a model in which the ladder is treated as a network of coupled small junctions arranged along both the edges and the rungs of the lad-
der. This model is often used to treat 2D Josephson networks, and includes no inductive coupling between junctions, other than that produced by the other junctions. To confirm our numerical results, we derive a discrete sine–Gordon equation from our coupled–network model. Thus, these seemingly quite different models produce nearly identical behavior for ladders. By extension, they suggest that some properties of 2D arrays might conceivably be treated by a similar simplification. In simulations [Bobbert (1992) Geigenmüller et al. (1993) Shea et al. (1995)], underdamped arrays of this type show some similarities to ladder arrays, exhibiting the analogs of both voltage steps and whirling modes.

We consider a ladder consisting of coupled superconducting grains, the $i^{th}$ of which has order parameter

$$\Phi_i = \Phi_0 e^{i\theta_i}.$$ 

Grains $i$ and $j$ are coupled by resistively–shunted Josephson junctions (RSJ’s) with current $I_{ij}$, shunt resistance $R_{ij}$ and shunt capacitance $C_{ij}$, with periodic boundary conditions.

The phases $\theta_i$ evolve according to the coupled RSJ equations

$$\hbar \dot{\theta}_i / (2e) = V_i,$n

$$M_{ij} \dot{V}_j = I_{ij}^{ext} / I_c - (R / R_{ij})(V_i - V_j) - (I_{ij} / I_c) \sin(\theta_{ij} - A_{ij}).$$

Here the time unit is $t_0 = \hbar / (2eRI_c)$, where $R$ and $I_c$ are the shunt resistance and critical current across a junction in the $x$–direction; $I_{ij}^{ext}$ is the external current fed into the $i^{th}$ node; the spatial distances are given in units of the lattice spacing $a$, and the voltage $V_i$ in units of $I_c R$.

$$M_{ij} = -4\pi eC I_c R^2 / \hbar$$

$$M_{ii} = -\sum_{j \neq i} M_{ij},$$

where $C$ is the intergrain capacitance. Finally,

$$A_{ij} = (2\pi / \Phi_0) \int_i^j A \cdot dl,$$

where $A$ is the vector potential. Following Ryu et al. (1996), we assume $N$ plaquettes in the array, and postulate a current $I$ uniformly injected into each node on the outer edge and extracted from each node on the inner edge of the ring. We also assume a uniform transverse magnetic field $B \equiv f R_0 / a^2$, and use the Landau gauge $A = -Bx \hat{y}$. 


We now show that this model reduces approximately to a discrete sine–Gordon equation for the phase differences. Label each grain by \((x,y)\) where \(x/a = 0, \ldots , N - 1\) and \(y/a = 0, 1\). Subtracting the equation of motion for \(\theta(x, 1)\) from that for \(\theta(x, 2)\), and defining

\[
\Psi(x) = \frac{1}{2} [\theta(x, 1) + \theta(x, 2)], \chi(x) = [\theta(x, 2) - \theta(x, 1)],
\]

we get a differential equation for \(\chi(x)\) which is second-order in time. This equation may be further simplified using the facts that \(A_{x,y;x\pm 1,y} = 0\) in the Landau gauge, and that \(A_{x,1;x,2} = -A_{x,2;x,1}\), and by defining the discrete Laplacian

\[
\chi(x + 1) - 2\chi(x) + \chi(x - 1) = \nabla^2 \chi(x).
\]

Finally, using the boundary conditions,

\[
I_{\text{ext}}(x, 2) = -I_{\text{ext}}(x, 1) \equiv I,
\]

and introducing \(\phi(x) = \chi(x) - A_{x,2;x,1}\), we get

\[
[1 - \eta_c^2 \nabla^2] \ddot{\phi} = i - [1 - \eta_c^2 \nabla^2] \dot{\phi} - \sin(\phi) + 2\eta_j^2 \sum_{i=\pm 1} \cos\{\Psi(x) - \Psi(x + i)\} \sin\{[\phi(x) - \phi(x + i)]/2\},
\]

where we have defined a dimensionless current \(i = I/I_{cy}\), and anisotropy factors

\[
2\eta_{c}^2 = R_y/R_x, \quad 2\eta_{j}^2 = C_x/C_y, \quad 2\eta_{j}^2 = I_{cx}/I_{cy}.
\]

We now neglect all combined space and time derivatives of order three or higher. Similarly, we set the cosine factor equal to unity (this is also checked numerically to be valid \textit{a posteriori}) and linearize the sine factor in the last term, so that the final summation can be expressed simply as \(\nabla^2 \phi\). With these approximations, \(6.87\) reduces to discrete driven sine-Gordon equation with dissipation:

\[
\beta \ddot{\phi} + \dot{\phi} + \sin(\phi) - \eta_c^2 \nabla^2 \phi = i, \quad \text{where} \quad \beta = 4\pi e I_{cy} R_y^2 C_y / h.
\]
Soliton Behavior

In the absence of damping and driving, the continuum version of (6.88) has, among other solutions, the sine–Gordon soliton [Rajaraman (1982)], given by

$$\phi_s(x, t) \sim 4 \tan^{-1}\left[\exp\left\{\frac{(x - v_v t)}{\sqrt{\eta_J^2 - \beta v_v^2}}\right\}\right]$$

where $v_v$ is the velocity. The phase in this soliton rises from $\sim 0$ to $\sim 2\pi$ in a width $d_k \sim \sqrt{\eta_J^2 - \beta v_v^2}$.

The transition to the resistive state occurs at $n_{\min} = 4, 2, 2, 1$ for $\eta_J^2 = 0, 5, 1, 25, 2, 5, 5$. This can also be understood from the kink–phason resonance picture. To a phason mode, the passage of a kink of width $d_k$ will appear like the switching on of a step–like driving current over a time of order $d_k/v_v$. The kink will couple to the phasons only if $d_k/v_v \geq \pi/\omega_1$, the half–period of the phason, or equivalently

$$\frac{1}{\sqrt{\beta v_v}} \geq \frac{\sqrt{1 + \pi^2}}{\eta_J} = 3.3 \frac{\eta}{\eta_J}.$$ 

This condition agrees very well with our numerical observations, even though it was obtained by considering soliton solutions from the continuum sine–Gordon equation.

The fact that the voltage in regime I is approximately linear in $f$ can be qualitatively understood from the following argument. Suppose that $\phi$ for $N_f$ fluxons can be approximated as a sum of well–separated solitons, each moving with the same velocity and described by

$$\phi(x, t) = \sum_{j=1}^{N_f} \phi_j,$$

where $\phi_j = \phi_s(x - x_j, t)$.

Since the solitons are well separated, we can use following properties:

$$\sin\left[\sum_j \phi_j\right] = \sum_j \sin \phi_j \quad \text{and} \quad \int \phi_j \dot{\phi}_j dx \propto \delta_{ij}.$$ 

By demanding that the energy dissipated by the damping of the moving soliton be balanced by that the driving current provides ($\propto \int dx \dot{\phi}(x)$), one can show that the $N_f$ fluxons should move with the same velocity $v$ as that for a single fluxon driven by the same current. In the whirling regime,
the $f-$independence of the voltage can be understood from a somewhat different argument. Here, we assume a periodic solution of the form
\[ \phi = \sum_{j}^{Nf} \phi_w(x - \tilde{v}t - j/f), \]
moving with an unknown velocity $\tilde{v}$ where $\phi_w(\xi)$ describes a whirling solution containing one fluxon. Then using the property $\phi(x + m/f) = \phi(x) + 2\pi m$, one can show that [Ryu et al. (1996)]
\[ \sin[\sum_{j}^{Nf} \phi_w(x - \tilde{v}t - j/f)] = \sin[Nf\phi_w(x - \tilde{v}t)]. \]

Finally, using the approximate property $\phi_w(\xi) \sim \xi$ of the whirling state, one finds $\tilde{v} = v/(Nf)$, leading to an $f-$independent voltage.

**Ballistic Soliton Motion and Soliton Mass**

A common feature of massive particles is their ‘ballistic motion’, defined as inertial propagation after the driving force has been turned off. Such propagation has been reported experimentally but as yet has not been observed numerically in either square or triangular lattices [Geigenmuller et al. (1993)]. In the so-called flux–flow regime at $\eta_J = 0.71$, we also find no ballistic propagation, presumably because of the large pinning energies produced by the periodic lattice.

We can define the fluxon mass in our ladder by equating the charging energy $E_c = C/2 \sum_{i,j} V_{ij}^2$ to the kinetic energy of a soliton of mass $M_v$:
\[ E_{kin} = \frac{1}{2} M_v v_v^2 \] [Geigenmuller et al. (1993)]. Since $E_c$ can be directly calculated in our simulation, while $v_v$ can be calculated from $\langle V \rangle$, this gives an unambiguous way to determine $M_v$. For $\eta_J = 0.5$, we find $E_c/C \sim 110(\langle V \rangle / I_c R)^2$, in the flux–flow regime. This gives $M_v \sim 3.4 C \phi_0^2/a^2$, more than six times the usual estimate for the vortex mass in a 2D square lattice. Similarly, the vortex friction coefficient $\gamma$ can be estimated by equating the rate of energy dissipation,
\[ E_{dis} = 1/2 \sum_{i,j} V_{ij}^2/R_{ij}, \quad \text{to} \quad \frac{1}{2} \gamma v_v^2. \]

This estimate yields $\gamma' \sim 3.4 \phi_0^2/(Ra^2)$, once again more than six times the value predicted for 2D arrays [Geigenmuller et al. (1993)]. This large dissipation explains the absence of ballistic motion for this anisotropy.
muller et al. (1993)]. At larger values \( \eta J = 5 \) and \( 2.5 \), a similar calculation gives \( M_1 \sim 0.28 \) and \( 0.34\phi_0^2/(Ra^2) \), \( \gamma^I \sim 0.28 \) and \( 0.34\phi_0^2/(Ra^2) \). These lower values of \( \gamma^I \), but especially the low pinning energies, may explain why ballistic motion is possible at these values of \( \eta J \). See [Ryu et al. (1996)] for more details.

### 6.5.4 Synchronization in Arrays of Josephson Junctions

The synchronization of coupled nonlinear oscillators has been a fertile area of research for decades [Pikovsky et al. (2001)]. In particular, Winfree–type phase models [Winfree (1967)] have been extensively studied. In 1D, a generic version of this model for \( N \) oscillators reads

\[
\dot{\theta}_j = \Omega_j + \sum_{k=1}^{N} \sigma_{j,k} \Gamma(\theta_k - \theta_j),
\]

where \( \theta_j \) is the phase of oscillator \( j \), which can be envisioned as a point moving around the unit circle with angular velocity \( \dot{\theta}_j = d\theta_j/dt \). In the absence of coupling, this overdamped oscillator has an angular velocity \( \Omega_j \). \( \Gamma(\theta_k - \theta_j) \) is the coupling function, and \( \sigma_{j,k} \) describes the range and nature (e.g., attractive or repulsive) of the coupling. The special case

\[
\Gamma(\theta_k - \theta_j) = \sin(\theta_k - \theta_j), \quad \sigma_{j,k} = \alpha/N, \quad \alpha = \text{const},
\]

corresponds to the uniform, sinusoidal coupling of each oscillator to the remaining \( N - 1 \) oscillators. This mean-field system is usually called the globally–coupled Kuramoto model (GKM). Kuramoto was the first to show that for this particular form of coupling and in the \( N \to \infty \) limit, there is a continuous dynamical phase transition at a critical value of the coupling strength \( \alpha_c \) and that for \( \alpha > \alpha_c \) both phase and frequency synchronization appear in the system [Kuramoto (1984); Strogatz (2000)]. If \( \sigma_{j,k} = \alpha \delta_{j,k+1} \) while the coupling function retains the form \( \Gamma(\theta_j - \theta_k) = \sin(\theta_k - \theta_j) \), then we have the so–called locally–coupled Kuramoto model (LKM), in which each oscillator is coupled only to its nearest neighbors. Studies of synchronization in the LKM [Sakaguchi (1987)], including extensions to more than one spatial dimension, have shown that \( \alpha_c \) grows without bound in the \( N \to \infty \) limit [Strogatz and Mirollo (1988)].

Watts and Strogatz introduced a simple model for tuning collections of coupled dynamical systems between the two extremes of random and regular networks [Watts and Strogatz (1998)]. In this model, connections
between nodes in a regular array are randomly rewired with a probability $p$, such that $p = 0$ means the network is regularly connected, while $p = 1$ results in a random connection of nodes. For a range of intermediate values of $p$ between these two extremes, the network retains a property of regular networks (a large clustering coefficient) and also acquires a property of random networks (a short characteristic path length between nodes). Networks in this intermediate configuration are termed small–world networks. Many examples of such small worlds, both natural and human–made, have been discussed [Strogatz (2001)]. Not surprisingly, there has been much interest in the synchronization of dynamical systems connected in a small–world geometry [Barahona and Pecora (2002)] [Nishikawa et al. (2003)].

Generically, such studies have shown that the presence of small–world connections make it easier for a network to synchronize, an effect generally attributed to the reduced path length between the linked systems. This has also been found to be true for the special case in which the dynamics of each oscillator is described by a Kuramoto model [Hong et al. (2002a)].

As an example of physically–controllable systems of nonlinear oscillators, which can be studied both theoretically and experimentally, Josephson junction (JJ) arrays are almost without peer. Through modern fabrication techniques and careful experimental methods one can attain a high degree of control over the dynamics of a JJ array, and many detailed aspects of array behavior have been studied [Newrock et al. (2000)]. Among the many different geometries of JJ arrays, ladder arrays deserve special attention. For example, they have been observed to support stable time–dependent, spatially–localized states known as discrete breathers [Tras et al. (2000)]. In addition, the ladder geometry is more complex than that of better understood serial arrays but less so than fully two–dimensional (2D) arrays. In fact, a ladder can be considered as a special kind of 2D array, and so the study of ladders could throw some light on the behavior of such 2D arrays. Also, linearly–stable synchronization of the horizontal, or rung, junctions in a ladder is observed in the absence of a load over a wide range of dc bias currents and junction parameters (such as junction capacitance), so that synchronization in this geometry appears to be robust [Trees et al. (2005)].

In the mid 1990’s it was shown that a serial array of zero-capacitance, i.e., overdamped, junctions coupled to a load could be mapped onto the GKM [Wiesenfeld et al. (1996)] [Wiesenfeld et al. (1998)]. The load in this case was essential in providing an all–to–all coupling among the junctions. The result was based on an averaging process, in which (at least)
two distinct time scales were identified: the ‘short’ time scale set by the
rapid voltage oscillations of the junctions (the array was current biased
above its critical current) and ‘long’ time scale over which the junctions
synchronize their voltages. If the resistively-shunted junction (RSJ) equa-
tions describing the dynamics of the junctions are integrated over one cycle
of the ‘short’ time scale, what remains is the ‘slow’ dynamics, describing
the synchronization of the array. This mapping is useful because it allows
knowledge about the GKM to be applied to understanding the dynamics of
the serial JJ array. For example, the authors of [Wiesenfeld et al. (1996)]
were able, based on the GKM, to predict the level of critical current dis-
order the array could tolerate before frequency synchronization would be
lost. Frequency synchronization, also described as entrainment, refers to
the state of the array in which all junctions not in the zero–voltage state
have equal (to within some numerical precision) time–averaged voltages:
\[ \langle \hbar/2e \rangle \langle \dot{\theta}_j \rangle, \]
where \( \theta_j \) is the gauge–invariant phase difference across junction
\( j \). More recently, the ‘slow’ synchronization dynamics of finite–capacitance
serial arrays of JJ’s has also been studied [Chernikov and Schmidt (1995);
Watanabe and Swift (1997)]. Perhaps surprisingly, however, no experi-
mental work on JJ arrays has verified the accuracy of this GKM mapping.
Instead, the first detailed experimental verification of Kuramoto’s theory
was recently performed on systems of coupled electrochemical oscillators
[Kiss et al. (2002)].

Recently, [Daniels et al. (2003)] showed, with an eye toward a better
understanding of synchronization in 2D JJ arrays, that a ladder array of
overdamped junctions could be mapped onto the LKM. This work was based
on an averaging process, as in [Wiesenfeld et al. (1996)], and was valid in
the limits of weak critical current disorder (less than about 10%) and large
dc bias currents, \( I_B \), along the rung junctions (\( I_B/\langle I_c \rangle \gtrsim 3 \), where \( \langle I_c \rangle \) is
the arithmetic average of the critical currents of the rung junctions. The
result demonstrated, for both open and periodic boundary conditions, that
synchronization of the current–biased rung junctions in the ladder is well
described by (6.89).

In this subsection, following [Trees et al. (2005)], we demonstrate that a
ladder array of underdamped junctions can be mapped onto a second–order
Winfree–type oscillator model of the form

\[
\ddot{\theta}_j + \dot{\theta}_j = \Omega_j + \sum_{k=1}^{N} \sigma_{j,k} \Gamma(\theta_k - \theta_j),
\]
where \( a \) is a constant related to the average capacitance of the rung junctions. This result is based on the \( \text{resistively \& capacitively–shunted junction (RCSJ)} \) model and a multiple time scale analysis of the classical equations for the array. Secondly, we study the effects of \( \text{small world (SW)} \) connections on the synchronization of both overdamped and underdamped ladder arrays. It appears that SW connections make it easier for the ladder to synchronize, and that a Kuramoto or Winfree type model (6.89) and (6.90), suitably generalized to include the new connections, accurately describes the synchronization of this ladder.

6.5.4.1 \( \text{Phase Model for Underdamped JJL} \)

Following \cite{Trees et al. (2005)} we analyze synchronization in disordered Josephson junction arrays. The ladder geometry used consists of an array with \( N = 8 \) plaquettes, periodic boundary conditions, and uniform dc bias currents, \( I_B \), along the rung junctions (see Figure 6.10). The \textit{gauge–invariant phase difference} across rung junction \( j \) is \( \gamma_j \), while the phase difference across the off–rung junctions along the outer(inner) edge of plaquette \( j \) is \( \psi_{1,j}(\psi_{2,j}) \). The critical current, resistance, and capacitance of rung junction \( j \) are denoted \( I_{cj}, R_j, \) and \( C_j \), respectively. For simplicity, we assume all off–rung junctions are identical, with critical current \( I_{co}, \) resistance \( R_o, \) and capacitance \( C_o \). We also assume that the product of the junction critical current and resistance is the same for all junctions in the array \cite{Benz (1995)}, with a similar assumption about the ratio of each junction’s critical current with its capacitance:

\[
I_{cj} R_j = I_{co} R_o = \frac{\langle I_c \rangle}{\langle R \rangle},
\]

\[
I_{cj} C_j = I_{co} C_o = \frac{\langle I_c \rangle}{\langle C \rangle},
\]

where for any generic quantity \( X \), the angular brackets with no subscript denote an arithmetic average over the set of rung junctions,

\[
\langle X \rangle \equiv \left( \frac{1}{N} \right) \sum_{j=1}^{N} X_j.
\]

For convenience, we work with dimensionless quantities. Our dimen-
Fig. 6.10  A ladder array of Josephson junctions with periodic boundary conditions and $N = 8$ plaquettes. A uniform, dc bias current $I_B$ is inserted into and extracted from each rung as shown. The gauge–invariant phase difference across the rung junctions is denoted by $\gamma_j$ where $1 \leq j \leq N$, while the corresponding quantities for the off–rung junctions along the outer(inner) edge are $\psi_{1,j}(\psi_{2,j})$ (see text for explanation).

The dimensionless time variable is

$$\tau \equiv \frac{t}{t_c} = \frac{2e\langle I_c \rangle t}{\hbar(R^{-1})}, \quad (6.93)$$

where $t$ is the ordinary time. In the following, derivatives with respect to $\tau$ will be denoted by prime (e.g., $\psi' = d\psi/d\tau$). The dimensionless bias current is

$$i_B \equiv \frac{I_B}{\langle I_c \rangle}, \quad (6.94)$$

while the dimensionless critical current of rung junction $j$ is $i_{cj} \equiv I_{cj}/\langle I_c \rangle$. The McCumber parameter in this case is

$$\beta_c \equiv \frac{2e\langle I_c \rangle \langle C \rangle}{\hbar(R^{-1})^2}, \quad (6.95)$$

Note that $\beta_c$ is proportional to the mean capacitance of the rung junctions.
An important dimensionless parameter is
\[ \alpha \equiv \frac{I_{co}}{\langle I_c \rangle}, \]  
(6.96)
which will effectively tune the nearest-neighbor interaction strength in our phase model for the ladder.

Conservation of charge applied to the superconducting islands on the outer and inner edge, respectively, of rung junction \( j \) yields the following equations in dimensionless variables [Trees et al. (2005)]:

\[
\begin{align*}
&i_B - ic_j \sin \gamma_j - ic_j \gamma_j' - ic_j \beta_c \gamma_j'' - \alpha \sin \psi_{1,j} - \alpha \psi_{1,j}' = 0, \\
&-\alpha \beta_c \psi_{1,j}' + \alpha \sin \psi_{1,j-1} + \alpha \psi_{1,j-1}' + \alpha \beta_c \psi_{1,j-1}'' = 0, \\
&-i_B + ic_j \sin \gamma_j + ic_j \gamma_j' + ic_j \beta_c \gamma_j'' - \alpha \sin \psi_{2,j} - \alpha \psi_{2,j}' = 0, \\
&-\alpha \beta_c \psi_{2,j}' + \alpha \sin \psi_{2,j-1} + \alpha \psi_{2,j-1}' + \alpha \beta_c \psi_{2,j-1}'' = 0,
\end{align*}
\]
(6.97)
where \( 1 \leq j \leq N \). The result is a set of \( 2N \) equations in \( 3N \) unknowns: \( \gamma_j \), \( \psi_{1,j} \), and \( \psi_{2,j} \). We supplement (6.98) by the constraint of fluxoid quantization in the absence of external or induced magnetic flux. For plaquette \( j \) this constraint yields the relationship
\[
\gamma_j + \psi_{2,j} - \gamma_{j+1} - \psi_{1,j} = 0.
\]
(6.99)

Equations (6.98) and (6.99) can be solved numerically for the \( 3N \) phases \( \gamma_j \), \( \psi_{1,j} \), and \( \psi_{2,j} \) [Trees et al. (2005)].

We assign the rung junction critical currents in one of two ways, randomly or nonrandomly. We generate random critical currents according to a parabolic probability distribution function (PDF) of the form
\[
P(i_c) = \frac{3}{4 \Delta^3} \left[ \Delta^2 - (i_c - 1)^2 \right],
\]
(6.100)
where \( i_c = I_c / \langle I_c \rangle \) represents a scaled critical current, and \( \Delta \) determines the spread of the critical currents. Equation (6.100) results in critical currents in the range \( 1 - \Delta \leq i_c \leq 1 + \Delta \). Note that this choice for the PDF (also used in [Wiesenfeld et al. (1996)]) avoids extreme critical currents (relative to a mean value of unity) that are occasionally generated by PDF’s with tails. The nonrandom method of assigning rung junction critical currents was based on the expression (with \( 1 \leq j \leq N \))
\[
i_{cj} = 1 + \Delta - \frac{2 \Delta}{(N-1)^2} \left[ 4j^2 - 4(N+1)j + (N+1)^2 \right],
\]
(6.101)
which results in the $i_{c,j}$ values varying quadratically as a function of position along the ladder and falling within the range $1 - \Delta \leq i_{c,j} \leq 1 + \Delta$. We usually use $\Delta = 0.05$.

**Multiple Time–Scale Analysis**

Now, our goal is to derive a Kuramoto–like model for the phase differences across the rung junctions, $\gamma_j$, starting with (6.98). We begin with two reasonable assumptions. First, we assume there is a simple phase relationship between the two off–rung junctions in the same plaquette [Trees et al. (2005)]:

$$\psi_{2,j} = -\psi_{1,j}$$

the validity of which has been discussed in detail in [Daniels et al. (2003); Filatrella and Wiesenfeld (1995)]. As a result, (6.99) reduces to

$$\psi_{1,j} = \frac{\gamma_j - \gamma_{j+1}}{2},$$

(6.103)

which implies that (6.97) can be written as

$$i_{c,j} \beta_c \gamma_j'' + i_{c,j} \gamma_j' + \frac{\alpha \beta_c}{2} [\gamma_j''_{j+1} - 2 \gamma_j'' + \gamma_j''_{j-1}] + \frac{\alpha}{2} [\gamma_j'_{j+1} - 2 \gamma_j' + \gamma_j'_{j-1}] = i_B - i_{c,j} \sin \gamma_j + \alpha \sum_{\delta = \pm 1} \sin \left(\frac{\gamma_{j+\delta} - \gamma_j}{2}\right).$$

(6.104)

Our second assumption is that we can neglect the discrete Laplacian terms in (6.104), namely

$$\nabla^2 \gamma_j' \equiv \gamma_{j+1}' - 2 \gamma_j' + \gamma_{j-1}'$$

and

$$\nabla^2 \gamma_j'' \equiv \gamma_{j+1}'' - 2 \gamma_j'' + \gamma_{j-1}''.$$

We find numerically, over a wide range of bias currents $i_B$, McCumber parameters $\beta_c$, and coupling strengths $\alpha$ that $\nabla^2 \gamma_j'$ and $\nabla^2 \gamma_j''$ oscillate with a time–averaged value of approximately zero. Since the multiple time scale method is similar to averaging over a fast time scale, it seems reasonable to drop these terms. In light of this assumption, (6.104) becomes

$$i_{c,j} \beta_c \gamma_j'' + i_{c,j} \gamma_j' = i_B - i_{c,j} \sin \gamma_j + \alpha \sum_{\delta = \pm 1} \sin \left(\frac{\gamma_{j+\delta} - \gamma_j}{2}\right).$$

(6.105)

We can use (6.105) as the starting point for a multiple time scale analysis. Following [Chernikov and Schmidt (1995)] and [Watanabe and Swift]
Quantum Leap (1997), we divide (6.105) by $i_B$ and define the following quantities:

\[ \tilde{\tau} \equiv i_B \tau, \quad \tilde{\beta}_c \equiv i_B \beta_c, \quad \epsilon = 1/i_B. \]  

(6.106)

In terms of these scaled quantities, (6.105) can be written as

\[ i_c j d^2 \gamma_j / d\tilde{\tau}^2 + i_c j d\gamma_j / d\tilde{\tau} + \epsilon i_c j \sin \gamma_j - \epsilon \alpha \sum_{\delta} \sin \left( \gamma_j + \delta - \gamma_j / 2 \right) = 1. \]  

(6.107)

Next, we introduce a series of four (dimensionless) time scales,

\[ T_n \equiv \epsilon^n \tilde{\tau}, \quad (n = 0, 1, 2, 3), \]  

(6.108)

which are assumed to be independent of each other. Note that $0 < \epsilon < 1$ since $\epsilon = 1/i_B$. We can think of each successive time scale, $T_n$, as being ‘slower’ than the scale before it. For example, $T_2$ describes a slower time scale than $T_1$. The time derivatives in (6.107) can be written in terms of the new time scales, since we can think of $\tilde{\tau}$ as being a function of the four independent $T_n$’s, $\tilde{\tau} = \tilde{\tau}(T_0, T_1, T_2, T_3)$. Letting $\partial_n \equiv \partial / \partial T_n$, the first and second time derivatives can be written as [Trees et al. (2005)]

\[ d / d\tilde{\tau} = \partial_0 + \epsilon \partial_1 + \epsilon^2 \partial_2 + \epsilon^3 \partial_3 \]  

(6.109)

\[ d^2 / d\tilde{\tau}^2 = \partial_0^2 + 2\epsilon \partial_0 \partial_1 + \epsilon^2 \left( 2\partial_0 \partial_2 + \partial_1^2 \right) + 2\epsilon^3 \left( \partial_0 \partial_3 + \partial_1 \partial_2 \right), \]  

(6.110)

where in (6.110) we have dropped terms of order $\epsilon^4$ and higher.

Next, we expand the phase differences in an $\epsilon$ expansion

\[ \gamma_j = \sum_{n=0}^{\infty} \epsilon^n \gamma_{n,j}(T_0, T_1, T_2, T_3). \]  

(6.111)

Substituting this expansion into (6.107) and collecting all terms of order $\epsilon^0$ results in the expression

\[ 1 = i_c j \tilde{\beta}_c \partial_0^2 \gamma_{0,j} + i_c j \partial_0 \gamma_{0,j}, \]  

(6.112)

for which we find the solution

\[ \gamma_{0,j} = \frac{T_0}{i_c j} + \phi_j(T_1, T_2, T_3), \]  

(6.113)

where we have ignored a transient term of the form $e^{-T_0 / \tilde{\beta}_c}$, and where $\phi_j(T_i)$, $(i = 1, 2, 3)$ is assumed constant over the fastest time scale $T_0$. Note
that the expression for $\gamma_{0,j}$ consists of a rapid phase rotation described by $T_0/i c_j$ and slower-scale temporal variations, described by $\phi_j$, on top of that overturning. In essence, the goal of this technique is to solve for the dynamical behavior of the slow phase variable, $\phi_j$. The resulting differential equation for the $\phi_j$ is [Trees et al. (2005)]:

$$
\beta_c \phi_j'' + \phi_j' = \Omega_j + K_j \sum_{\delta = \pm 1} \sin \left( \frac{\phi_{j+\delta} - \phi_j}{2} \right) + L_j \sum_{\delta = \pm 1} \sin \left( 3 \left( \frac{\phi_{j+\delta} - \phi_j}{2} \right) \right) \\
+ M_j \sum_{\delta = \pm 1} \left\{ \cos \left( \frac{\phi_{j+\delta} - \phi_j}{2} \right) - \cos \left( 3 \left( \frac{\phi_{j+\delta} - \phi_j}{2} \right) \right) \right\},
$$

(6.114)

where $\Omega_j$ is given by the expression (letting $x_j \equiv i c_j/i B$ for convenience)

$$
\Omega_j = \frac{1}{x_j} \left[ 1 - \frac{x_j^4}{(2\beta_c^2 + x_j^2)^2} \right],
$$

(6.115)

and the three coupling strengths are

$$
K_j = \frac{\alpha i c_j}{i c_j} \left[ 1 + \frac{x_j^4 \left( 3x_j^2 + 23\beta_c^2 \right)}{16 \left( \beta_c^2 + x_j^2 \right)^2} \right],
$$

(6.116)

$$
L_j = \frac{\alpha x_j^4 (3\beta_c^2 - x_j^2)}{i c_j 16 \left( \beta_c^2 + x_j^2 \right)^2},
$$

(6.117)

$$
M_j = -\frac{\alpha x_j^5 \beta_c}{i c_j 4 \left( \beta_c^2 + x_j^2 \right)^2}.
$$

(6.118)

We emphasize that (6.114) is expressed in terms of the original, unscaled, time variable $\tau$ and McCumber parameter $\beta_c$.

We will generally consider bias current and junction capacitance values such that $x_j^2 \ll \beta_c^2$. In this limit, (6.116)–(6.118) can be approximated as follows [Trees et al. (2005)]:

$$
K_j \rightarrow \frac{\alpha}{i c_j} \left[ 1 + \mathcal{O} \left( \frac{1}{x_B^4} \right) \right],
$$

(6.119)

$$
L_j \rightarrow \frac{\alpha x_j^4 (3\beta_c^2 - x_j^2)}{16 \beta_c^2} \sim \mathcal{O} \left( \frac{1}{i B^4} \right),
$$

(6.120)

$$
M_j \rightarrow -\frac{\alpha x_j^5 \beta_c}{4 \beta_c^2} \sim \mathcal{O} \left( \frac{1}{i B^4} \right).
$$

(6.121)
For large bias currents, it is reasonable to truncate Equation (6.114) at $O(1/i^3 B)$, which leaves

$$\beta_c \phi''_j + \phi'_j = \Omega_j + \alpha \sum_{k \neq j, \delta = \pm 1} \sin \left[ \frac{\phi_{j+k} - \phi_j}{2} \right],$$

(6.122)

where all the cosine coupling terms and the third harmonic sine term have been dropped as a result of the truncation.

In the absence of any coupling between neighboring rung junctions ($\alpha = 0$) the solution to (6.122) is

$$\phi_j^{(\alpha=0)} = A + Be^{-\tau/\beta_c} + \Omega_j \tau,$$

where $A$ and $B$ are arbitrary constants. Ignoring the transient exponential term, we see that $d\phi_j^{(\alpha=0)}/d\tau = \Omega_j$, so we can think of $\Omega_j$ as the voltage across rung junction $j$ in the un–coupled limit. Alternatively, $\Omega_j$ can be viewed as the angular velocity of the strongly–driven rotator in the un–coupled limit.

Equation (6.122) is our desired phase model for the rung junctions of the underdamped ladder [Trees et al. (2005)]. The result can be described as a locally–coupled Kuramoto model with a second-order time derivative (LKM2) and with junction coupling determined by $\alpha$. In the context of systems of coupled rotators, the second derivative term is due to the non–negligible rotator inertia, whereas in the case of Josephson junctions the second derivative arises because of the junction capacitance. The globally–coupled version of the second–order Kuramoto model (GKM2) has been well studied; in this case the oscillator inertia leads to a first–order synchronization phase transition as well as to hysteresis between a weakly and a strongly coherent synchronized state [Tanaka et al. (1997b); Acebrón et al. (2000)].

6.5.4.2 Comparison of LKM2 and RCSJ Models

We now compare the synchronization behavior of the RCSJ ladder array with the LKM2. We consider frequency and phase synchronization separately. For the rung junctions of the ladder, frequency synchronization occurs when the time average voltages, $\langle v_j \rangle_\tau = \langle \phi'_j \rangle_\tau$ are equal for all $N$ junctions, within some specified precision. In the language of coupled rotators, this corresponds to phase points moving around the unit circle with the same average angular velocity. We quantify the degree of frequency
synchronization via an ‘order parameter’ [Trees et al. (2005)]

\[
f = 1 - \frac{s_v(\alpha)}{s_v(0)}, \tag{6.123}
\]

where \( s_v(\alpha) \) is the standard deviation of the \( N \) time–average voltages, \( \langle v_j \rangle_\tau \):

\[
s_v(\alpha) = \sqrt{\frac{\sum_{j=1}^{N} \left( \langle v_j \rangle_\tau - \frac{1}{N} \sum_{k=1}^{N} \langle v_k \rangle_\tau \right)^2}{N-1}} \tag{6.124}
\]

In general, this standard deviation will be a function of the coupling strength \( \alpha \), so \( s_v(0) \) is a measure of the spread of the \( \langle v_j \rangle_\tau \) values for \( N \) independent junctions. Frequency synchronization of all \( N \) junctions is signaled by \( f = 1 \), while \( f = 0 \) means all \( N \) average voltages have their un–coupled values.

Phase synchronization of the rung junctions is measured by the usual Kuramoto order parameter

\[
r \equiv \frac{1}{N} \sum_{j=1}^{N} e^{i\phi_j}. \tag{6.125}
\]

Lastly in this subsection, we address the issue of the linear stability of the frequency synchronized states (\( \alpha > \alpha_c \)) by calculating their Floquet exponents numerically for the RCSJ model as well as analytically based on the LKM2, \((6.122)\). The analytic technique used has been described in [Trees and Murgescu (2001)], giving as a result for the real part of the Floquet exponents:

\[
\text{Re}(\lambda_m t_c) = -\frac{1}{2\beta_c} \left[ 1 \pm \text{Re} \sqrt{1 - 4\beta_c (K + 3L) \omega_m^2} \right], \tag{6.126}
\]

where stable solutions correspond to exponents, \( \lambda_m \), with a negative real part. One can think of the \( \omega_m \) as the normal mode frequencies of the ladder. We find that for a ladder with periodic boundary conditions and \( N \) plaquettes

\[
\omega_m^2 = \frac{4 \sin^2 \left( \frac{m\pi}{N} \right)}{1 + 2 \sin^2 \left( \frac{m\pi}{N} \right)}, \quad 0 \leq m \leq N - 1. \tag{6.127}
\]

To arrive at \((6.126)\) we have ignored the effects of disorder so that \( K \) and \( L \) are obtained from \((6.116)\) and \((6.117)\) with the substitution \( i_{cj} \rightarrow 1 \) throughout. This should be reasonable for the levels of disorder we have
considered (5%). Substituting the expressions for \( \bar{K} \) and \( \bar{L} \) into (6.126) results in \[ \text{Trees et al. (2005)} \]

\[
\text{Re}(\lambda m_{tc}) = -\frac{1}{2\beta_c} \left[ 1 \pm \text{Re} \sqrt{1 - 2\beta_c \alpha \left( 1 + \frac{2\beta_c^2}{(\gamma \beta_c^2 + 1)^2} \right)} \right] \omega_m^2. \quad (6.128)
\]

We are most interested in the Floquet exponent of minimum magnitude, \( \text{Re}(\lambda_{\min} m_{tc}) \), which essentially gives the lifetime of the longest-lived perturbations to the synchronized state.

6.5.4.3 ‘Small–World’ Connections in J JL Arrays

Many properties of small world networks have been studied in the last several years, including not only the effects of network topology but also the dynamics of the node elements comprising the network \[ \text{Newman (2000)} \]
\[ \text{Strogatz (2001)} \]. Of particular interest has been the ability of oscillators to synchronize when configured in a small–world manner. Such synchronization studies can be broadly sorted into several categories \[ \text{Trees et al. (2005)} \]:

(1) Work on coupled lattice maps has demonstrated that synchronization is made easier by the presence of random, long–range connections \[ \text{Gade and Hu (2000)} \]
\[ \text{Batista et al. (2003)} \].

(2) Much attention has been given to the synchronization of continuous time dynamical systems, including the first order locally–coupled Kuramoto model (LKM), in the presence of small–world connections \[ \text{Hong et al. (2002a)} \]
\[ \text{Hong et al. (2002b)} \]
\[ \text{Watts (1999)} \]. For example, Hong and coworkers \[ \text{Hong et al. (2002a)} \]
\[ \text{Hong et al. (2002b)} \] have shown that the LKM, which does not exhibit a true dynamical phase transition in the thermodynamic limit \( (N \to \infty) \) in the pristine case, does exhibit such a phase synchronization transition for even a small number of shortcuts.

But the assertion \[ \text{Wang and Chen (2002)} \] that any small world network can synchronize for a given coupling strength and large enough number of nodes, even when the pristine network would not synchronize under the same conditions, is not fully accepted \[ \text{Barahona and Pecora (2002)} \].

(3) More general studies of synchronization in small world and scale–free networks \[ \text{Barahona and Pecora (2002)} \]
\[ \text{Nishikawa et al. (2003)} \] have shown that the small world topology does not guarantee that a network can synchronize. In \[ \text{Barahona and Pecora (2002)} \] it was shown that one could calculate the average number of shortcuts per node, \( s_{\text{sync}} \), required for a
given dynamical system to synchronize. This study found no clear relation between this synchronization threshold and the onset of the small world region, i.e., the value of $s$ such that the average path length between all pairs of nodes in the array is less than some threshold value. Nishikawa et al. (2003) studied arrays with a power–law distribution of node connectivities (scale–free networks) and found that a broader distribution of connectivities makes a network less synchronizable even though the average path length is smaller. It was argued that this behavior was caused by an increased number of connections on the hubs of the scale–free network. Clearly it is dangerous to assume that merely reducing the average path length between nodes of an array will make such an array easier to synchronize.

Now, regarding Josephson–junction arrays, if we have a disordered array biased such that some subset of the junctions are in the voltage state, i.e., undergoing limit cycle oscillations, the question is will the addition of random, long–range connections between junctions aid the array in attaining frequency and/or phase synchronization? Can we address this question by using the mapping discussed above between the RCSJ model for the underdamped ladder array and the second–order, locally–coupled Kuramoto model (LKM2). Based on the results of Daniels et al. (2003), we also know that the RSJ model for an overdamped ladder can be mapped onto a first–order, locally–coupled Kuramoto model (LKM). Because of this mapping, the ladder array falls into category (2) of the previous paragraph. In other words, we should expect the existence of shortcuts to drastically improve the ability of ladder arrays to synchronize Trees et al. (2005).

We add connections between pairs of rung junctions that will result in interactions that are longer than nearest neighbor in range. We do so by adding two, nondisordered, off–rung junctions for each such connection. We argue that the RCSJ equations for the underdamped junctions in the ladder array can be mapped onto a straightforward variation of (6.122), in which the sinusoidal coupling term for rung junction $j$ also includes the longer–range couplings due to the added shortcuts. Imagine a ladder with a shortcut between junctions $j$ and $l$, where $l \neq j, j \pm 1$. Conservation of charge applied to the two superconducting islands that comprise rung junction $j$ will lead to equations very similar to (6.98). For example, the analog to (6.97) will be

\[
i_B - i_{ej} \sin \gamma_j - i_{ej}' \gamma_j' - \beta \epsilon i_{ej} \gamma_j'' - \alpha \sin \psi_{1,j} - \alpha \psi_{j,1}' \\
- \beta \epsilon \alpha \psi_{j,1}'' + \alpha \sin \psi_{1,j-1} + \alpha \psi_{1,j-1}' + \beta \epsilon \alpha \psi_{1,j-1}'' \\
+ \sum_l \left[ \alpha \sin \psi_{1,j;l} + \alpha \psi_{1,j;l}' + \beta \epsilon \alpha \psi_{1,j;l}'' \right] = 0,
\]
with an analogous equation corresponding to the inner superconducting island that can be generalized from (6.98). The sum over the index \( l \) accounts for all junctions connected to junction \( j \) via an added shortcut. Fluxoid quantization still holds, which means that we can augment (6.99) with

\[
\gamma_j + \psi_{2;jl} - \gamma_l - \psi_{1;jl} = 0. \tag{6.129}
\]

We also assume the analog of (6.102) holds:

\[
\psi_{2;jl} = -\psi_{1;jl}. \tag{6.130}
\]

Equations (6.129) and (6.130) allow us to write the analog to (6.103) for the case of shortcut junctions:

\[
\psi_{1;jl} = \frac{\gamma_j - \gamma_l}{2}. \tag{6.131}
\]

Equation (6.129), in light of (6.131), can be written as

\[
i_B - i_{ej} \sin \gamma_j - i_{ej} \gamma_j' - \beta_c i_{ej} \gamma_j'' + \alpha \sum_{d=\pm 1} \sin \left( \frac{\gamma_j + d - \gamma_j}{2} \right) + \alpha \sum_l \sin \left( \frac{\gamma_j - \gamma_l}{2} \right) + \frac{\alpha}{2} \nabla^2 \gamma_j' + \frac{\alpha}{2} \nabla^2 \gamma_j'' + \frac{\alpha}{2} \sum_l (\gamma_j' - \gamma_l') + \frac{\alpha}{2} \sum_l (\gamma_j'' - \gamma_l'') = 0,
\]

where the sums \( \Sigma_i \) are over all rung junctions connected to \( j \) via an added shortcut. As we did with the pristine ladder, we will drop the two discrete Laplacians, since they have a very small time average compared to the terms \( i_{ej} \gamma_j' + i_{ej} \beta_c \gamma_j'' \). The same is also true, however, of the terms \( \alpha / 2 \sum_l (\gamma_j' - \gamma_l') \) and \( \alpha / 2 \sum_l (\gamma_j'' - \gamma_l'') \), as direct numerical solution of the full RCSJ equations in the presence of shortcuts demonstrates. So we shall drop these terms as well. Then we have

\[
i_B - i_{ej} \sin \gamma_j - i_{ej} \gamma_j' - \beta_c i_{ej} \gamma_j'' + \frac{\alpha}{2} \sum_{k \in \Lambda_j} \sin \left( \frac{\gamma_k - \gamma_j}{2} \right), \tag{6.132}
\]

where the sum is over all junctions in \( \Lambda_j \), which is the set of all junctions connected to junction \( j \). From above results we can predict that a multiple time scale analysis of (6.132) results in a phase model of the form

\[
\beta_c \frac{d^2 \phi_j}{d\tau^2} + \frac{d\phi_j}{d\tau} = \Omega_j + \frac{\alpha}{2} \sum_{k \in \Lambda_j} \sin \left( \frac{\phi_k - \phi_j}{2} \right), \tag{6.133}
\]
where $\Omega_j$ is given by (6.115). A similar analysis for the overdamped ladder leads to the result
\[ \phi_j' = \Omega_j^{(1)} + \frac{\alpha}{2} \sum_{k \in \Lambda_j} \sin \left( \phi_k - \phi_j \right), \] (6.134)
where the time–averaged voltage across each overdamped rung junction in the un–coupled limit is
\[ \Omega_j^{(1)} = \sqrt{\left( \frac{iB}{i_{c3}} \right)^2 - 1}. \] (6.135)

Although the addition of shortcuts makes it easier for the array to synchronize, we should also consider the effects of such random connections on the stability of the synchronized state. The Floquet exponents for the synchronized state allow us to quantify this stability. Using a general technique discussed in [Pecora and Carroll (1998)], we can calculate the Floquet exponents $\lambda_m$ for the LKM based on the expression
\[ \lambda_m t_c = \alpha E_m^G, \] (6.136)
where $E_m^G$ are the eigenvalues of $G$, the matrix of coupling coefficients for the array. A specific element, $G_{ij}$, of this matrix is unity if there is a connection between rung junctions $i$ and $j$. The diagonal terms, $G_{ii}$, is merely the negative of the number of junctions connected to junction $i$. This gives the matrix the property $\sum_j G_{ij} = 0$. In the case of the pristine ladder, the eigenvalues of $G$ can be calculated analytically, which yields Floquet exponents of the form
\[ \lambda_m^{(p=0)} t_c = -4\alpha \sin^2 \left( \frac{m\pi}{N} \right). \] (6.137)

See [Trees et al. (2005)] for more details.
Chapter 7

Quantum Mind

In this chapter we present a quantum theory of the mind and consciousness.

7.1 Quantum Theory of Consciousness

7.1.1 Human Mind

Recall that the word mind commonly refers to the collective aspects of intellect and consciousness which are manifest in some combination of thought, perception, emotion, will, memory, and imagination.

There are many theories of what the mind is and how it works, dating back to Plato, Aristotle and other Ancient Greek philosophers. Modern theories, based on a scientific understanding of the brain, see the mind as a phenomenon of psychology, and the term is often used more or less synonymously with consciousness.

The question of which human attributes make up the mind is also much debated. Some argue that only the ‘higher’ intellectual functions constitute mind: particularly reason and memory. In this view the emotions – love, hate, fear, joy – are more ‘primitive’ or subjective in nature and should be seen as different in nature or origin to the mind. Others argue that the rational and the emotional sides of the human person cannot be separated, that they are of the same nature and origin, and that they should all be considered as part of the individual mind.

In popular usage mind is frequently synonymous with thought: It is that private conversation with ourselves that we carry on ‘inside our heads’ during every waking moment of our lives. Thus we ‘make up our minds,’ or ‘change our minds’ or are ‘of two minds’ about something. One of the key attributes of the mind in this sense is that it is a private sphere. No–
else can ‘know our mind.’ They can only know what we communicate.

Both philosophers and psychologists remain divided about the nature of the mind. Some take what is known as the substantial view, and argue that the mind is a single entity, perhaps having its base in the brain but distinct from it and having an autonomous existence. This view ultimately derives from Plato, and was absorbed from him into Christian thought. In its most extreme form, the substantial view merges with the theological view that the mind is an entity wholly separate from the body, in fact a manifestation of the soul, which will survive the body’s death and return to God, its creator.

Others take what is known as the functional view, ultimately derived from Aristotle, which holds that the mind is a term of convenience for a variety of mental functions which have little in common except that humans are conscious of their existence. Functionalists tend to argue that the attributes which we collectively call the mind are closely related to the functions of the brain and can have no autonomous existence beyond the brain, nor can they survive its death. In this view mind is a subjective manifestation of consciousness: the human brain’s ability to be aware of its own existence. The concept of the mind is therefore a means by which the conscious brain understands its own operations.

A leading exponent of the substantial view at the mind was George Berkeley, an 18th century Anglican bishop and philosopher. Berkeley argued that there is no such thing as matter and what humans see as the material world is nothing but an idea in God’s mind, and that therefore the human mind is purely a manifestation of the soul or spirit. This type of belief is also common in certain types of spiritual non-dualistic belief, but outside this field few philosophers take an extreme view today. However, the view that the human mind is of a nature or essence somehow different from, and higher than, the mere operations of the brain, continues to be widely held.

Berkeley’s views were attacked, and in the eyes of many philosophers demolished, by T.H. Huxley, a 19th century biologist and disciple of Charles Darwin.
Darwin, who agreed that the phenomena of the mind were of a unique order, but argued that they can only be explained in reference to events in the brain. Huxley drew on a tradition of materialist thought in British philosophy dating to Thomas Hobbes, who argued in the 17th century that mental events were ultimately physical in nature, although with the biological knowledge of his day he could not say what their physical basis was. Huxley blended Hobbes with Darwin to produce the modern functional view. Huxley’s view was reinforced by the steady expansion of knowledge about the functions of the human brain. In the 19th century it was not possible to say with certainty how the brain carried out such functions as memory, emotion, perception and reason. This left the field open for substantialists to argue for an autonomous mind, or for a metaphysical theory of the mind. But each advance in the study of the brain during the 20th century made this harder, since it became more and more apparent that all the components of the mind have their origins in the functioning of the brain. Huxley’s rationalism, was disturbed in the early 20th century by Freudian a theory of the unconscious mind, and argued that those mental processes of which humans are subjectively aware are only a small part of their total mental activity.

---

2Charles Robert Darwin (12 February 1809 – 19 April 1882) was an English naturalist who achieved lasting fame by producing considerable evidence that species originated through evolutionary change, at the same time proposing the scientific theory that natural selection is the mechanism by which such change occurs. This theory is now considered a cornerstone of biology.

3Thomas Hobbes (April 5, 1588–December 4, 1679) was an English philosopher, whose famous 1651 book Leviathan set the agenda for nearly all subsequent Western political philosophy. Although Hobbes is today best remembered for his work on political philosophy, he contributed to a diverse array of fields, including history, geometry, ethics, general philosophy and what would now be called political science. Additionally, Hobbes’s account of human nature as self-interested cooperation has proved to be an enduring theory in the field of philosophical anthropology.
More recently, Douglas Hofstadter's 1979 Pulitzer Prize–winning book 'Gödel, Escher, Bach – an eternal Golden Braid', is a tour de force on the subject of mind, and how it might arise from the neurology of the brain. Amongst other biological and cybernetic phenomena, Hofstadter places tangled loops and recursion at the center of self, self-awareness, and perception of oneself, and thus at the heart of mind and thinking. Likewise philosopher Ken Wilber posits that Mind is the interior dimension of the brain holon, i.e., mind is what a brain looks like internally, when it looks at itself.

Quantum physicist David Bohm had a theory of mind that is most

4Douglas Richard Hofstadter (born February 15, 1945 in New York, New York) is an American academic, the son of Nobel Prize–winning physicist Robert Hofstadter. He is probably best known for his book Gödel, Escher, Bach: an Eternal Golden Braid (abbreviated as GEB) which was published in 1979, and won the 1980 Pulitzer Prize for general non-fiction. This book is commonly considered to have inspired many students to begin careers in computing and artificial intelligence, and attracted substantial notice outside its central artificial intelligence readership owing to its drawing on themes from such diverse disciplines as high-energy physics, music, the visual arts, molecular biology, and literature.

5David Joseph Bohm (born December 20, 1917 in Wilkes–Barre, Pennsylvania, died October 27, 1992 in London) was an American–born quantum physicist, who made significant contributions in the fields of theoretical physics, philosophy and neuropsychology, and to the Manhattan Project.

Bohm made a number of significant contributions to physics, particularly in the area of quantum mechanics and relativity theory. While still a post-graduate at Berkeley, he developed a theory of plasmas, discovering the electron phenomenon now known as Bohm–diffusion. His first book, Quantum Theory published in 1951, was well–received by Einstein, among others. However, Bohm became dissatisfied with the orthodox approach to quantum theory, which he had written about in that book, and began to develop his own approach (Bohm interpretation), a non–local hidden variable deterministic theory whose predictions agree perfectly with the nondeterministic quantum theory. His work and the EPR argument became the major factor motivating John Bell's inequality, whose consequences are still being investigated.

Bohm also made significant theoretical contributions to neuropsychology and the development of the so-called holonomic brain model. In collaboration with Stanford neuroscientist Karl Pribram, Bohm helped establish the foundation for Pribram’s theory that the brain operates in a manner similar to a hologram, in accordance with quantum mathematical principles and the characteristics of wave patterns. These wave forms may compose hologram–like organizations, Bohm suggested, basing this concept on his application of Fourier analysis, a mathematical method for decomposing complex waves into component sine waves. The holonomic brain model developed by Pribram and Bohm posits a lens defined world view, much like the textured prismatic effect of sunlight refracted by the churning mists of a rainbow, a view which is quite different from the more conventional 'objective' approach. Pribram believes that if psychology means to understand the conditions that produce the world of appearances, it must look to the thinking of physicists like Bohm.

Bohm proposes thus in his book ‘Thought as a System’ a pervasive, systematic nature
comparable to Neo–Platonic theories. “Thought runs you. Thought, however, gives false info that you are running it, that you are the one who controls thought. Whereas actually thought is the one which controls each one of us...” [Bohm (1992)].

The debate about the nature of the mind is relevant to the development of artificial intelligence (see next section). If the mind is indeed a thing separate from or higher than the functioning of the brain, then presumably it will not be possible for any machine, no matter how sophisticated, to duplicate it. If on the other hand the mind is no more than the aggregated functions of the brain, then it will be possible, at least in theory, to create a machine with a mind.

Currently, the Mind/Brain/Behavior Interfaculty Initiative (MBB) at Harvard University aims to elucidate the structure, function, evolution, development, and pathology of the nervous system in relation to human behavior and mental life. It draws on the departments of psychology, neurobiology, neurology, molecular and cellular biology, radiology, psychiatry, organismic and evolutionary biology, history of science, and linguistics.

On the other hand, human brain has been considered (by E.M. Izhikevich, Editor of the new Encyclopedia of Computational Neuroscience) as a weakly–connected neural network, with possibly chaotic behavior [Izhikevich (1999b)], consisting of $n$ quasi–periodic cortical oscillators $X_1, ..., X_n$ forced by the thalamic input $X_0$ (see Figure 7.1).
7.1.1.1 The Mind–Body Problem

The mind–body problem is essentially the problem of explaining the relationship between minds, or mental processes, and bodily states or processes (see, e.g., [Kim (1995a)]). Our perceptual experiences depend on stimuli which arrive at our various sensory organs from the external world and that these stimuli cause changes in the states of our brain, ultimately causing us to feel a sensation which may be pleasant or unpleasant. Someone’s desire for a slice of pizza will tend to cause that person to move their body in a certain manner in a certain direction in an effort to get what they want. But how is it possible that conscious experiences can arise out of an inert lump of gray matter endowed with electrochemical properties? [Kim (1995b)]. How does someone’s desire cause that individual’s neurons to fire and his muscles to contract in exactly the right manner? These are some of the essential puzzles that have confronted philosophers of mind at least from the time of René Descartes.

6René Descartes (March 31, 1596 – February 11, 1650), also known as Cartesius, was a noted French philosopher, mathematician, and scientist. Dubbed the ‘Founder of Modern Philosophy’ and the ‘Father of Modern Mathematics’, he ranks as one of the most important and influential thinkers of modern times. Much of subsequent western philosophy is a reaction to his writings, which have been closely studied from his time down to the present day. Descartes was one of the key thinkers of the Scientific Revolution in the Western World. He is also honoured by having the Cartesian coordinate system used in plane geometry and algebra named after him.

Descartes was a major figure in 17th century continental rationalism, later advocated by Baruch Spinoza and Gottfried Leibniz, and opposed by the empiricist school.
Dualism

Recall that dualism is a set of views about the relationship between mind and matter, which begins with the claim that mental phenomena are, in some respects, non–physical [Hartle (1996)]. One of the earliest known formulations of mind–body dualism existed in the eastern Sankhya school of Hindu philosophy (c. 650 BCE) which divided the world into Purusha (mind/spirit) and Prakrti (material substance). In the Western philosophical tradition, we first encounter similar ideas with the writings of Plato and Aristotle, who maintained, for different reasons, that man’s intelligence could not be identified with, or explained in terms of, his physical body (see, e.g., [Robinson (1983)]). However, the best–known version of dualism is due to René Descartes (1641), and holds that the mind is a non–physical substance [Descartes (1991)]. Descartes was the first to clearly identify the mind with consciousness and self–awareness and to distinguish this from the brain, which was the seat of intelligence. Hence, he was the first to formulate the mind–body problem in the form in which it still exists today.

The main argument in favor of dualism is simply that it appeals to the common–sense intuition of the vast majority of non–philosophically–trained people. If asked what the mind is, the average person will usually respond by identifying it with their self, their personality, their soul, or some other such entity, and they will almost certainly deny that the mind simply is the brain or vice–versa, finding the idea that there is just one ontological entity at play to be too mechanistic or simply unintelligible [Hartle (1996)]. The majority of modern philosophers of mind reject dualism, suggesting that these intuitions, like many others, are probably misleading. We should use our critical faculties, as well as empirical evidence from the sciences, to examine these assumptions and determine if there is any real basis to them [Hartle (1996)]. Another very important, more modern, argument in favor of dualism consists in the idea that the mental and the physical seem to have quite different and perhaps irreconcilable properties [Jackson (1982)].
Mental events have a certain subjective quality to them, whereas physical events obviously do not. For example, what does a burned finger feel like? What does blue sky look like? What does nice music sound like? Philosophers of mind call the subjective aspects of mental events qualia (or raw feels) [Jackson (1982)]. There is something that it is like to feel pain, to see a familiar shade of blue, and so on; there are qualia involved in these mental events. And the claim is that qualia seem particularly difficult to reduce to anything physical [Nagel (1974)].

Interactionist dualism, or simply interactionism, is the particular form of dualism first espoused by Descartes in the ‘Meditations’ [Descartes (1991)]. In the 20th century, its major defenders have been Karl Popper and John Eccles (see Popper and Eccles (2002)). It is the view that

7Sir Karl Raimund Popper (July 28, 1902 – September 17, 1994), was an Austrian and British philosopher and a professor at the London School of Economics. He is counted among the most influential philosophers of science of the 20th century, and also wrote extensively on social and political philosophy. Popper is perhaps best known for repudiating the classical observationalist–inductivist account of scientific method by advancing empirical falsifiability as the criterion for distinguishing scientific theory from non-science; and for his vigorous defense of liberal democracy and the principles of social criticism which he took to make the flourishing of the ‘open society’ possible. In 1934 he published his first book, ‘The Logic of Scientific Discovery’, in which he criticized psychologism, naturalism, inductionism, and logical positivism, and put forth his theory of potential falsifiability being the criterion for what should be considered science.

Popper coined the term critical rationalism to describe his philosophy. This designation is significant, and indicates his rejection of classical empiricism, and of the observationalist-inductivist account of science that had grown out of it. Popper argued strongly against the latter, holding that scientific theories are universal in nature, and can be tested only indirectly, by reference to their implications. He also held that scientific theory, and human knowledge generally, is irreducibly conjectural or hypothetical, and is generated by the creative imagination in order to solve problems that have arisen in specific historico-cultural settings. Logically, no number of positive outcomes at the level of experimental testing can confirm a scientific theory, but a single genuine counterexample is logically decisive: it shows the theory, from which the implication is derived, to be false. Popper’s account of the logical asymmetry between verification and falsification lies at the heart of his philosophy of science. This inspired him to take falsifiability as his criterion of demarcation between what is and is not genuinely scientific: a theory should be considered scientific if and only if it is falsifiable. This led him to attack the claims of both psychoanalysis and contemporary Marxism to scientific status, on the basis that the theories enshrined by them are not falsifiable. His scientific work was influenced by his study of quantum mechanics (he has written extensively against the famous Copenhagen interpretation) and by Albert Einstein’s approach to scientific theories.

In his book ‘All Life is Problem Solving’ (1999), Popper sought to explain the apparent progress of scientific knowledge, how it is that our understanding of the universe seems to improve over time. This problem arises from his position that the truth content of our theories, even the best of them, cannot be verified by scientific testing, but can only
mental states, such as beliefs and desires, causally interact with physical states \cite{Hartle(1996)}. Descartes’ famous argument for this position can be summarized as follows: Fred has a clear and distinct idea of his mind as a thinking thing which has no spatial extension (i.e., it cannot be measured in terms of length, weight, height, and so on) and he also has a clear and distinct idea of his body as something that is spatially extended, subject to quantification and not able to think. It follows that mind and body are

be falsified. If so, then how is it that the growth of science appears to result in a growth in knowledge? In Popper’s view, the advance of scientific knowledge is an evolutionary process characterised by his formula:

$$PS_1 \rightarrow TT_1 \rightarrow EE_1 \rightarrow PS_2.$$ 

In response to a given problem situation, $PS_1$, a number of competing conjectures, or tentative theories, $TT_1$, are systematically subjected to the most rigorous attempts at falsification possible. This process, error elimination, $EE_1$, performs a similar function for science that natural selection performs for biological evolution. Theories that better survive the process of refutation are not more true, but rather, more ‘fit’, in other words, more applicable to the problem situation at hand, $PS_1$. Consequently, just as a species ‘biological fit’ does not predict continued survival, neither does rigorous testing protect a scientific theory from refutation in the future. Yet, as it appears that the engine of biological evolution has produced, over time, adaptive traits equipped to deal with more and more complex problems of survival, likewise, the evolution of theories through the scientific method may, in Popper’s view, reflect a certain type of progress: toward more and more interesting problems, $PS_2$. For Popper, it is in the interplay between the tentative theories (conjectures) and error elimination (refutation) that scientific knowledge advances toward greater and greater problems; in a process very much akin to the interplay between genetic variation and natural selection.

As early as 1934 Popper wrote of the search for truth as one of the “strongest motives for scientific discovery.” Still, he describes in ‘Objective Knowledge’ (1972) early concerns about the much–criticised notion of truth as correspondence. Then came the semantic theory of truth formulated by the logician Alfred Tarski. Popper writes of learning in 1935 of the consequences of Tarski’s theory, to his intense joy. The theory met critical objections to truth as correspondence and thereby rehabilitated it. The theory also seemed to Popper to support metaphysical realism and the regulative idea of a search for truth.

Among his contributions to philosophy is his answer to David Hume’s ‘Problem of Induction’. Hume stated that just because the sun has risen every day for as long as anyone can remember, doesn’t mean that there is any rational reason to believe it will come up tomorrow. There is no rational way to prove that a pattern will continue on just because it has before. Popper’s reply is characteristic, and ties in with his criterion of falsifiability. He states that while there is no way to prove that the sun will come up, we can theorize that it will. If it does not come up, then it will be disproven, but since right now it seems to be consistent with our theory, the theory is not disproven. Thus, Popper’s demarcation between science and non–science serves as an answer to an old logical problem as well. This approach was criticised by Peter Singer for masking the role induction plays in empirical discovery.
not identical because they have radically different properties, according to Descartes [Descartes (1991)]. At the same time, however, it is clear that Fred’s mental states (desires, beliefs, etc.) have causal effects on his body and vice-versa: a child touches a hot stove (physical event) which causes pain (mental event) and makes him yell (physical event) which provokes a sense of fear and protectiveness in the mother (mental event) and so on. Descartes’ argument obviously depends on the crucial premise that what Fred believes to be ‘clear and distinct’ ideas in his mind are necessarily true. Most modern philosophers doubt the validity of such an assumption, since it has been shown in modern times by Freud (a third–person psychologically–trained observer can understand a person’s unconscious motivations better than she does), by Pierre Duhem (a third–person philosopher of science can know a person’s methods of discovery better than she does), by

8Sir John Carew Eccles (January 27, 1903 – May 2, 1997) was an Australian neurophysiologist who won the 1963 Nobel Prize in Physiology or Medicine for his work on the synapse. He shared the prize together with Andrew Fielding Huxley and Alan Lloyd Hodgkin.

In the early 1950s, Eccles and his colleagues performed the key experiments that would win Eccles the Nobel Prize. To study synapses in the peripheral nervous system, Eccles and colleagues used the stretch reflex as a model. This reflex is easily studied because it consists of only two neurons: a sensory neuron (the muscle spindle fiber) and the motor neuron. The sensory neuron synapses onto the motor neuron in the spinal cord. When Eccles passed a current into the sensory neuron in the quadriceps, the motor neuron innervating the quadriceps produced a small excitatory postsynaptic potential (EPSP). When he passed the same current through the hamstring, the opposing muscle to the quadriceps, he saw an inhibitory postsynaptic potential (IPSP) in the quadriceps motor neuron. Although a single EPSP was not enough to fire an action potential in the motor neuron, the sum of several EPSPs from multiple sensory neurons synapsing onto the motor neuron could cause the motor neuron to fire, thus contracting the quadriceps. On the other hand, IPSPs could subtract from this sum of EPSPs, preventing the motor neuron from firing.

Apart from these seminal experiments, Eccles was key to a number of important developments in neuroscience. Until around 1949, Eccles believed that synaptic transmission was primarily electrical rather than chemical. Although he was wrong in this hypothesis, his arguments led himself and others to perform some of the experiments which proved chemical synaptic transmission. Bernard Katz and Eccles worked together on some of the experiments which elucidated the role of acetylcholine as a neurotransmitter.

9Pierre Maurice Marie Duhem (10 June 1861 – 14 September 1916) French physicist and philosopher of science. Duhem’s sophisticated views on the philosophy of science are explicated in ‘The aim and structure of physical theory’ (foreword by Prince Louis de Broglie). In this work he refuted the inductivist untruth that Newton’s laws can be deduced from Kepler, et al. (a selection was published as Medieval cosmology: theories of infinity, place, time, void, and the plurality of worlds. He gave his name to the Quine-Duhem thesis, which holds that for any given set of observations there are an innumerably large number of explanations. Thus empirical evidence cannot force the revision of a theory.
Bronislaw Malinowski\textsuperscript{10} (an anthropologist can know a person’s customs and habits better than he does), and by theorists of perception (experiments can make one see things that are not there and scientists can describe a person’s perceptions better than he can), that such an idea of privileged and perfect access to one’s own ideas is dubious at best.

Other important forms of dualism which arose as reactions to, or attempts to salvage, the Cartesian version are:

(i) Psycho–physical parallelism, or simply parallelism, is the view that mind and body, while having distinct ontological statuses, do not causally influence one another, but run along parallel paths (mind events causally interact with mind events and brain events causally interact with brain events) and only seem to influence each other \cite{Robinson1983}. This view was most prominently defended by Gottfried Leibniz\textsuperscript{11}. Although Leibniz was actually an ontological monist who believed that only one fundamental substance, monads, exists in the universe and everything else is reducible to it, he nonetheless maintained that there was an important distinction between ‘the mental’ and ‘the physical’ in terms of causation. He held that God had arranged things in advance so that minds and bodies would be in harmony with each other. This is known as the doctrine of pre–established harmony \cite{Leibniz1714}.

(ii) Occasionalism is the view espoused by Nicholas Malebranche which asserts that all supposedly causal relations between physical events or be-

\textsuperscript{10}Bronislaw Kasper Malinowski (April 7, 1884 – May 16, 1942) was a Polish anthropologist widely considered to be one of the most important anthropologists of the twentieth century because of his pioneering work on ethnographic fieldwork, the study of reciprocity, and his detailed contribution to the study of Melanesia.

\textsuperscript{11}Gottfried Wilhelm Leibniz (July 1 (June 21 Old Style) 1646 – November 14, 1716) was a German polymath. Educated in law and philosophy, Leibniz played a major role in the European politics and diplomacy of his day. He occupies an equally large place in both the history of philosophy and the history of mathematics. He invented calculus independently of Newton, and his notation is the one in general use since. He also invented the binary system, foundation of virtually all modern computer architectures. In philosophy, he is most remembered for optimism, i.e., his conclusion that our universe is, in a restricted sense, the best possible one God could have made. He was, along with René Descartes and Baruch Spinoza, one of the three great 17th century rationalists, but his philosophy also both looks back to the Scholastic tradition and anticipates logic and analysis. Leibniz also made major contributions to physics and technology, and anticipated notions that surfaced much later in biology, medicine, geology, probability theory, psychology, knowledge engineering, and information science. He also wrote on politics, law, ethics, theology, history, and philology, even occasional verse. His contributions to this vast array of subjects are scattered in journals and in tens of thousands of letters and unpublished manuscripts. To date, there is no complete edition of Leibniz’s writings, and a complete account of his accomplishments is not yet possible.
tween physical and mental events are not really causal at all. While body and mind are still different substances on this view, causes (whether mental or physical) are related to their effects by an act of God’s intervention on each specific occasion [Schmaltz (2002)].

(iii) Epiphenomenalism is a doctrine first formulated by Thomas Huxley [Huxley (1898)]. Fundamentally, it consists in the view that mental phenomena are causally inefficacious. Physical events can cause other physical events and physical events can cause mental events, but mental events cannot cause anything, since they are just causally inert by-products (i.e. epiphenomena) of the physical world [Robinson (1983)] The view has been defended most strongly in recent times by Frank Jackson [Jackson (1982)].

(iv) Property dualism asserts that when matter is organized in the appropriate way (i.e., in the way that living human bodies are organized), mental properties emerge. Hence, it is a sub-branch of emergent materialism [Hartle (1996)]. These emergent properties have an independent ontological status and cannot be reduced to, or explained in terms of, the physical substrate from which they emerge. This position is espoused by David Chalmers and has undergone something of a renaissance in recent years [Chalmers (1997)].

Monism

In contrast to dualism, monism states that there is only one fundamental substance. Monism, first proposed in the West by Parmenides and in modern times by Baruch Spinoza[12] maintains that there is only one substance; in the East, rough parallels might be the Hindu concept of Brahman, or the Tao of Lao Tzu [Spinoza (1670)]. Today the most common

[12]Parmenides of Elea (early 5th century BC) was an ancient Greek philosopher born in Elea, a Hellenic city on the southern coast of Italy. Parmenides was a student of Ameinias and the founder of the School of Elea, which also included Zeno of Elea and Melissus of Samos.

[13]Benedictus de Spinoza (November 24, 1632 – February 21, 1677), named Baruch Spinoza by his synagogue elders, was a Jewish–Dutch philosopher. He is considered one of the great rationalists of 17th-century philosophy and, by virtue of his magnum opus the ‘Ethics’, one of the definitive ethicists. His writings, like those of his fellow rationalists, reveal considerable mathematical training and facility. Spinoza was a lens crafter by trade, an exciting engineering field at the time because of great discoveries being made by telescopes. The full impact of his work only took effect some time after his death and after the publication of his ‘Opera Posthuma’. He is now seen as having prepared the way for the 18th century Enlightenment, and as a founder of modern biblical criticism. 20th century philosopher, Gilles Deleuze (1990), referred to Spinoza as “The absolute philosopher, whose Ethics is the foremost book on concepts.”
forms of monism in Western philosophy are physicalistic [Kim (1995b)]. Physicalistic monism asserts that the only existing substance is physical, in some sense of that term to be clarified by our best science [Stoljar (2005)]. Another form of monism is that which states that the only existing substance is mental. Such idealistic monism is currently somewhat uncommon in the West [Kim (1995b)].

Phenomenalism, the theory that all that exists are the representations (or sense data) of external objects in our minds and not the objects themselves, was adopted by Bertrand Russell[14] and many of the logical posi-

[14]Bertrand Arthur William Russell, (3rd Earl Russell, 18 May 1872 – 2 February 1970), was a British philosopher, logician, and mathematician, working mostly in the 20th century. A prolific writer, Bertrand Russell was also a popularizer of philosophy and a commentator on a large variety of topics, ranging from very serious issues to the mundane. Continuing a family tradition in political affairs, he was a prominent liberal as well as a socialist and anti-war activist for most of his long life. Millions looked up to Russell as a prophet of the creative and rational life; at the same time, his stances on many topics were extremely controversial.

Russell was born at the height of Britain’s economic and political ascendancy. He died of influenza nearly a century later, at a time when the British Empire had all but vanished, its power dissipated by two debilitating world wars. As one of the world’s best-known intellectuals, Russell’s voice carried great moral authority, even into his early 90s. Among his political activities, Russell was a vigorous proponent of nuclear disarmament and an outspoken critic of the American war in Vietnam.

In 1950, Russell was made a Nobel Laureate in Literature, “in recognition of his varied and significant writings in which he champions humanitarian ideals and freedom of thought.”

Russell is generally recognized as one of the founders of analytical philosophy, even of its several branches. At the beginning of the 20th century, alongside G.E. Moore, Russell was largely responsible for the British ‘revolt against Idealism’, a philosophy greatly influenced by Georg Hegel. This revolt was echoed 30 years later in Vienna by the logical positivists’ ‘revolt against metaphysics’. Russell was particularly appalled by the idealist doctrine of internal relations, which held that in order to know any particular thing, we must know all of its relations. Russell showed that this would make space, time, science and the concept of number unintelligible. Russell’s logical work with Alfred Whitehead continued this project.

Russell had great influence on modern mathematical logic. His first mathematical book, An Essay on the Foundations of Geometry, was published in 1897. This work was heavily influenced by Immanuel Kant. Russell soon realised that the conception it laid out would have made Albert Einstein’s schema of space-time impossible, which he understood to be superior to his own system. Thenceforth, he rejected the entire Kantian program as it related to mathematics and geometry, and he maintained that his own earliest work on the subject was nearly without value. Russell discovered that Gottlob Frege had independently arrived at equivalent definitions for 0, successor, and number, and the definition of number is now usually referred to as the Frege–Russell definition. It was largely Russell who brought Frege to the attention of the English-speaking world. He did this in 1903, when he published ‘The Principles of Mathematics’, in which the concept of class is inextricably tied to the definition of number. The appendix to this work
activists during the early 20th century [Russell (1918)]. It lasted for only a very brief period of time. A third possibility is to accept the existence of a basic substance which is neither physical nor mental. The mental and physical would both be properties of this neutral substance. Such a position was adopted by Baruch Spinoza [Spinoza (1670)] and popularized by Ernst Mach [Mach (1959)] in the 19th century. This neutral monism, as it is called, resembles property dualism.

The Brahman of the Upanishads

Closely related to monism is the fundamental Vedantic concept of the Brahman. Sri Chinmoy writes [Sri Chinmoy (1974)]: “The heart of the Upanishads is most meaningful and most fruitful because it embodies the Life of the Brahman. Brahman is Reality in existence; Brahman is Reality’s existence. The eternal Truth of the Brahman is in the finite, beyond the finite, in the Infinite, and beyond the ever-transcending Infinite.

In the domain of realisation, Brahman is the Sovereign Absolute. In the domain of revelation, Brahman is the Omnipresent Reality. And in detailed a paradox arising in Frege’s application of second- and higher-order functions which took first-order functions as their arguments, and he offered his first effort to resolve what would henceforth come to be known as the Russell Paradox, which he later developed into a complete theory, the Theory of types. Aside from exposing a major inconsistency in naive set theory, Russell’s work led directly to the creation of modern axiomatic set theory. It also crippled Frege’s project of reducing arithmetic to logic. The Theory of Types and much of Russell’s subsequent work have also found practical applications with computer science and information technology.

Russell continued to defend logicism, the view that mathematics is in some important sense reducible to logic, and along with his former teacher, Alfred Whitehead, wrote the monumental Principia Mathematica, an axiomatic system on which all of mathematics can be built. The first volume of the Principia was published in 1910, and is largely ascribed to Russell. More than any other single work, it established the specialty of mathematical or symbolic logic. Two more volumes were published, but their original plan to incorporate geometry in a fourth volume was never realised, and Russell never felt up to improving the original works, though he referenced new developments and problems in his preface to the second edition. Upon completing the Principia, three volumes of extraordinarily abstract and complex reasoning, Russell was exhausted, and he never felt his intellectual faculties fully recovered from the effort. Although the Principia did not fall prey to the paradoxes in Frege’s approach, it was later proven by Kurt Gödel that neither Principia Mathematica, nor any other consistent system of primitive recursive arithmetic, could, within that system, determine that every proposition that could be formulated within that system was decidable, i.e., could decide whether that proposition or its negation was provable within the system (Gödel’s incompleteness theorem).

15Ernst Mach (February 18, 1838 – February 19, 1916) was an Austrian–Czech physicist and philosopher and is the namesake for the ‘Mach number’ (aka Mach speed) and the optical illusion known as Mach bands.
the field of manifestation, Brahman is the immortalizing Perfection. Brah-
man the Creator is the Consciousness–Light; Brahman the Fulfiller is the
Consciousness–Delight. Brahman is the inner Soul of all and the only Goal
in all.

When we look within, Brahman is Consciousness–Force. When we look
without, Brahman is Self–Manifestation. When we think of Brahman with
the mind, earth–bound mind, limited mind, sophisticated mind, unaspiring
mid, our life becomes sheer frustration. But when we meditate on Brah-
man in the heart, in the silent recesses of the heart, our life becomes pure
illumination.

To a non–seeker, Brahman is unknowable. To a beginner–seeker, Brah-
man is unknown. To a master–seeker, Brahman is knowable, Brahman
is known. Further, he himself grows into the Consciousness of Brahman.
Sarvam khalu idam Brahma, Indeed, all is Brahman.

The Eternal is existence within. The Eternal is existence without. There
is no abiding happiness in the finite. It is only in the Infinite that we
can hear the message of eternal Delight: Anandam Brahma and Anantam
Brahma. These are the two major aspects of Brahman. Anandam Brahma
is the life of the all–illuminating Delight and the all–fulfilling Delight. Anan-
tam Brahma is the Life of Infinity.

Here on earth the Life of Infinity constantly grows for the fulfilment of
the Absolute Brahman. That is why the Upanishadic Seers sing from the
depths of their hearts about the transcendental Delight of the Brahman:
Anandadd he eva khalv imani bhutani jayante,
Anandena jatani jivanti,
Anandam prayantyabhisam visanti.
From the transcendental Delight we came into existence; in Delight we grow
and play our respective roles; and at the end of our journeys close we enter
into the Supreme Delight.

Again, when the Seers saw Infinity in Brahman, they sang:
Aum, purnam adah, purnam idam, purnat purnam udacyate
Purnasya purnam adaya purnam evavatsisyate.
Infinity is that. Infinity is this. From Infinity, Infinity has come into exis-
tence. From Infinity, when Infinity is taken away, Infinity remains.

Brahman is active. Brahman is inactive. The active Brahman inwardly
does and outwardly becomes. Also, the active Brahman outwardly does
and inwardly becomes. But the inactive Brahman is the total Freedom of
inaction and complete Freedom in inaction.

The whole universe came into existence from Brahman the Seed. When
Brahman wanted to project Himself, He just projected Himself through four significant worlds: Ambhas, the highest world; Marichi, the sky; Mara the mortal world, the earth; and Apa, the world beneath earth. Then Brahman sent forth the guardians of these worlds. Next, He sent forth food for them. Then Brahman came to realise that He Himself had to take part in His Cosmic Game, so He entered into the Cosmic Lila (Game) through His own Yogic power. First He entered into the human body through the skull. The door by which Brahman entered is called the Door of Delight. This door is the highest center of consciousness. This is known as Sahasrara, the thousand–petaled lotus. It is situated in the center of the brain. The realisation of the Yogi enters there and becomes one with the Consciousness of the Brahman.

Brahman has many names, but His secret name is AUM.
Pranavo dhamuh saro atma ... AUM is the bow and Atma, the Self, is the arrow; Brahman is the target.

Through repeated practice the arrow is fixed into the target, the Brahmic Consciousness. That is to say, through regular concentration, meditation, and contemplation, the seeker enters into the Absolute Consciousness of the Brahman.

Creation is the supreme sacrifice of the Brahman. Creation is by no means a mechanical construction. Creation is a spiritual act, supremely revealing, manifesting, and fulfilling the divine splendor of the Brahman. The divine Architect is beyond creation, and at the same time manifests Himself in and through creation.

Brahman created out of His Being priests, warriors, tradesmen, and servants.
Then He created the Law. Nothing can be higher than this Law. This Law is Truth. When a man speaks the Truth, he declares the Law. When he declares the Law, he speaks the Truth. The Truth and the Law are one, inseparable.

Indian mythology has divided Time – not earthbound time but eternal Time – into four divisions: satya yuga, treta yuga, dwapara yuga, and kali yuga. According to many we are now in the kali yuga. Brahman in the kali yuga is fast asleep. He is in inconscience–ignorance–mire. In the dwapara yuga He awakes and He looks around. In the treta yuga He stands up, about to move forward. In the satya yuga, the Golden age, He moves fast, faster, fastest, towards His Goal. The message of the Vedas, the eternal message of the Aryan culture and civilization, the realisation of the Indian Sages and Seers, is movement, inner progress, the lifes march towards the
destined Goal.
Charai veti, charai veti. Move on, move on.”

**Behaviorism**

Behaviorism dominated philosophy of mind for much of the 20th century, especially the first half [Kim (1995b)]. In psychology, *behaviorism* developed as a reaction to the inadequacies of introspectionism. Introspective reports on one’s own interior mental life are not subject to careful examination for accuracy and are not generalizable. Without generalizability and the possibility of third-person examination, the behaviorists argued, science is simply not possible [Stoljar (2005)]. The way out for psychology was to eliminate the idea of an interior mental life (and hence an ontologically independent mind) altogether and focus instead on the description of observable behavior [Skinner (1972)].

Parallel to these developments in psychology, a philosophical behaviorism (sometimes called logical behaviorism) was developed [Stoljar (2005)]. This is characterized by a strong verificationism, which generally considers unverifiable statements about interior mental life senseless. But what are mental states if they are not interior states on which one can make introspective reports? The answer of the behaviorist is that mental states do not exist but are actually just descriptions of behavior and/or dispositions to behave made by external third parties in order to explain and predict others’ behavior [Ryle (1949)]. Philosophical behaviorism is considered by most modern philosophers of mind to be outdated [Kim (1995a)]. Apart from other problems, behaviorism implausibly maintains, for example, that someone is talking about behavior if she reports that she has a wracking headache.

**Continental Philosophy of Mind**

In contrast to Anglo–American *analytic philosophy*\(^\text{16}\) there are other schools of thought which are sometimes subsumed under the broad label

\(^{16}\) Analytic philosophy is the dominant academic philosophical movement in English–speaking countries and in the Nordic countries. It is distinguished from Continental Philosophy which pertains to most non–English speaking countries. Its main founders were the Cambridge philosophers G.E. Moore and Bertrand Russell. However, both were heavily influenced by the German philosopher and mathematician Gottlob Frege and many of analytic philosophy’s leading proponents, such as Ludwig Wittgenstein, Rudolf Carnap, Kurt Gödel, Karl Popper, Hans Reichenbach, Herbert Feigl, Otto Neurath, and Carl Hempel have come from Germany and Austria. In Britain, Russell and Moore were succeeded by C. D. Broad, L. Stebbing, Gilbert Ryle, A. J. Ayer, R. B. Braithwaite,
of continental philosophy. These schools tend to differ from the analytic school in that they focus less on language and logical analysis and more on directly understanding human existence and experience. With reference specifically to the discussion of the mind, this tends to translate into attempts to grasp the concepts of thought and perceptual experience in some direct sense that does not involve the analysis of linguistic forms [Dummett (2001)]. In particular, in his ‘Phenomenology of Mind’, G.W. F. Hegel.\textsuperscript{17}

Paul Grice, John Wisdom, R. M. Hare, J. L. Austin, P. F. Strawson, William Kneale, G. E. M. Anscombe, and Peter Geach. In America, the movement was led by many of the above-named European emigres as well as Max Black, Ernest Nagel, C. L. Stevenson, Norman Malcolm, W. V. Quine, Wilfrid Sellars, and Nelson Goodman, while A. N. Prior, John Passmore, and J. J. C. Smart were prominent in Australasia.

Logic and philosophy of language were central strands of analytic philosophy from the beginning, although this dominance has diminished greatly. Several lines of thought originate from the early, language-and-logic part of this analytic philosophy tradition. These include: logical positivism, logical empiricism, logical atomism, logicism and ordinary language philosophy. Subsequent analytic philosophy includes extensive work in ethics (such as Philippa Foot, R. M. Hare, and J. L. Mackie), political philosophy (John Rawls, Robert Nozick, Richard Montague, Hilary Putnam, W.V.O. Quine, Nathan Salmon, John Searle), and philosophy of mind (Daniel Dennett, David Chalmers, Putnam). Analytic metaphysics has also recently come into its own (Kripke, David Lewis, Salmon, Peter van Inwagen, P.F. Strawson).

\textsuperscript{17}Georg Wilhelm Friedrich Hegel (August 27, 1770 – November 14, 1831) was a German philosopher born in Stuttgart, Württemberg, in present-day southwest Germany. His influence has been widespread on writers of widely varying positions, including both his admirers (F.H. Bradley, J.P. Sartre, Hans Küng, Bruno Bauer), and his detractors (Kierkegaard, Schopenhauer, Heidegger, Schelling). His great achievement was to introduce for the first time in philosophy the idea that History and the concrete are important in getting out of the circle of philosophia perennis, i.e., the perennial problems of philosophy. Also, for the first time in the history of philosophy he realised the importance of the Other in the coming to be of self-consciousness, see slave–master dialectic.

Some of Hegel’s writing was intended for those with advanced knowledge of philosophy, although his ‘Encyclopedia’ was intended as a textbook in a university course. Nevertheless, like many philosophers, Hegel assumed that his readers would be well-versed in Western philosophy, up to and including Descartes, Spinoza, Hume, Kant, Fichte, and Schelling. For those wishing to read his work without this background, introductions to Hegel and commentaries about Hegel may suffice. However, even this is hotly debated since the reader must choose from multiple interpretations of Hegel’s writings from incompatible schools of philosophy. Presumably, reading Hegel directly would be the best method of understanding him, but this task has historically proved to be beyond the average reader of philosophy.[citation needed] This difficulty may be the most urgent problem with respect to the legacy of Hegel.

One especially difficult aspect of Hegel’s work is his innovation in logic. In response to Immanuel Kant’s challenge to the limits of Pure Reason, Hegel developed a radically new form of logic, which he called speculation, and which is today popularly called dialectics. The difficulty in reading Hegel was perceived in Hegel’s own day, and persists
discusses three distinct types of mind: the subjective mind, the mind of an individual; the objective mind, the mind of society and of the State; and the Absolute mind, a unity of all concepts. In modern times, the two main schools that have developed in response or opposition to this Hegelian tradition are phenomenology and existentialism. Phenomenology, founded by Edmund Husserl, focuses on the contents of the human mind and how phenomenological processes shape our experiences. Existentialism, a school of thought led by Jean–Paul Sartre, focuses on the content of experiences into the 21st century. To understand Hegel fully requires paying attention to his critique of standard logic, such as the law of contradiction and the law of the excluded middle, and, whether one accepts or rejects it, at least taking it seriously. Many philosophers who came after Hegel and were influenced by him, whether adopting or rejecting his ideas, did so without fully absorbing his new speculative or dialectical logic.

Edmund Gustav Albrecht Husserl (April 8, 1859, Prostějov – April 26, 1938, Freiburg) was a German philosopher, known as the father of phenomenology. Husserl was born into a Jewish family in Prostějov (Prossnitz), Moravia, Czech Republic (then part of the Austrian Empire). A pupil of Franz Brentano and Carl Stumpf, Husserl came to influence, among others, Edith Stein (St. Teresa Benedicta of the Cross), Eugen Fink, Martin Heidegger, Jean–Paul Sartre, and Maurice Merleau–Ponty; in addition, Hermann Weyl's interest in intuitionistic logic and impredicativity appear to have resulted from contacts with Husserl. Rudolf Carnap was also influenced by Husserl, not only concerning Husserl’s notion of essential insight that Carnap used in his Der Raum, but also his notion of formation rules and transformation rules is founded on Husserl's philosophy of logic. In 1887 Husserl converted to Christianity and joined the Lutheran Church. He taught philosophy at Halle as a tutor (Privatdozent) from 1887, then at Göttingen as professor from 1901, and at Freiburg im Breisgau from 1916 until he retired in 1928. After this, he continued his research and writing by using the library at Freiburg, until barred therefrom because of his Jewish heritage under the rectorship of his former pupil and intended protege, Martin Heidegger.

Husserl held the belief that truth-in-itself has as ontological correlate being-in-itself, just as meaning categories have formal-ontological categories as correlates. The discipline of logic is a formal theory of judgment, that studies the formal a priori relations among judgments using meaning categories. Mathematics, on the other hand, is formal ontology, it studies all the possible forms of being (of objects). So, in both of these disciplines, formal categories, in their different forms, are their object of study, not the sensible objects themselves. The problem with the psychological approach to mathematics and logic is that it fails to account for the fact that it is about formal categories, not abstractions from sensibility alone. The reason why we do not deal with sensible objects in mathematics is because of another faculty of understanding called categorial abstraction. Through this faculty we are able to get rid of sensible components of judgments, and just focus on formal categories themselves. Thanks to ‘eidetic (or essential) intuition’, we are able to grasp the possibility, impossibility, necessity and contingency among concepts or among formal categories. Categorial intuition, along with categorial abstraction and eidetic intuition, are the basis for logical and mathematical knowledge.

Jean–Paul Charles Aymard Sartre (June 21, 1905 – April 15, 1980), was a French existentialist philosopher, dramatist and screenwriter, novelist and critic.

The basis of Sartre’s existentialism is found in his ‘The Transcendence of the Ego’. To begin with, the thing-in-itself is infinite and overflowing. Any direct consciousness
and how the mind deals with such experiences [Flynn (2004)].

Neurobiology

On the other hand, within the tangible field of neurobiology, there are many subdisciplines which are concerned with the relations between mental and physical states and processes [Bear et al. (1995)]:

(1) Sensory neurophysiology investigates the relation between the processes of perception and stimulation [Pinel (1997)].
(2) Cognitive neuroscience studies the correlations between mental processes and neural processes [Pinel (1997)].
(3) Neuropsychology describes the dependence of mental faculties on specific anatomical regions of the brain [Pinel (1997)].
(4) Lastly, evolutionary biology studies the origins and development of the human nervous system and, in as much as this is the basis of the mind, also describes the ontogenetic and phylogenetic development of mental phenomena beginning from their most primitive stages [Pinker (1997)].

Since the 1980’s, sophisticated neuroimaging procedures, such as fMRI, have furnished increasing knowledge about the workings of the human brain, shedding light on ancient philosophical problems. The methodological breakthroughs of the neurosciences, in particular the introduction of high-tech neuroimaging procedures, has propelled scientists toward the elaboration of increasingly ambitious research programs: one of the main goals is to describe and comprehend the neural processes which correspond to mental functions [Bear et al. (1995)]. A very small number of neurobiologists, such as Emil Reymond and John Eccles have denied the possibility of a thing–in–itself, Sartre refers to as a ‘pre–reflective consciousness’. Any attempt to describe, understand, historicize etc. the thing–in–itself, Sartre calls ‘reflective consciousness’. There is no way for the reflective consciousness to subsume the pre–reflective, and so reflection is fated to a form of anxiety, i.e., the human condition. The reflective consciousness in all its forms, (scientific, artistic or otherwise) can only limit the thing–in–itself by virtue of its attempt to understand or describe it. It follows therefore that any attempt at self-knowledge (self-consciousness) is a construct that fails no matter how often it is attempted. (self-consciousness is a reflective consciousness of an overflowing infinite) In Sartre’s words “Consciousness is consciousness of itself insofar as it is consciousness of a transcendent object.” The same holds true about knowledge of the ‘Other’ (being), which is a construct of reflective consciousness. One must be careful to understand this more as a form of warning than as an ontological statement. However, there is an implication of Solipsism here that Sartre considers fundamental to any coherent description of the human condition.

20Emil du Bois–Reymond (November 7, 1818, Berlin, Germany – November 26, 1896),
‘reduction’ of mental phenomena to cerebral processes (see [1]). However, the contemporary neurobiologist and philosopher Gerhard Roth continues to defend a form of ‘non-reductive materialism’ [2].

7.1.2 EPR Paradox and Bell’s Theorem

Now we move to quantum realm. We start with EPR Paradox and Bell’s Theorem.

7.1.2.1 EPR Paradox

Recall that the so-called EPR paradox in quantum mechanics is a thought experiment which challenged long-held ideas about the relation between the observed values of physical quantities and the values that can be accounted for by a physical theory. ‘EPR’ stands for A. Einstein, was a German physician and physiologist, discoverer of the nerve action potential and the father of experimental electrophysiology. ‘EPR’ stands for A. Einstein, was a German–born theoretical physicist. He is widely regarded as one of the greatest physicists who ever lived. He formulated the Special and General Relativity Theories. In addition, he made significant contributions to quantum theory and statistical mechanics. While best known for the Theory of Relativity (and specifically mass-energy equivalence, $E = mc^2$), he was awarded the 1921 Nobel Prize for Physics for his explanation of the photoelectric effect in 1905 (his ‘wonderful year’ or ‘miraculous year’) and ‘for his services to Theoretical Physics’.

Following the May 1919 British solar-eclipse expeditions, whose later analysis confirmed that light rays from distant stars were deflected by the Sun’s gravitation as predicted by the Field Equation of general relativity, in November 1919 Albert Einstein became world-famous, an unusual achievement for a scientist. The Times ran the headline on November 7, 1919: “Revolution in science – New theory of the Universe – Newtonian ideas overthrown.” Nobel laureate Max Born viewed General Relativity as the “greatest feat of human thinking about nature;” fellow laureate Paul Dirac called...
B. Podolsky, and N. Rosen, who introduced the thought experiment in a 1935 paper [Einstein et al. (1935a)], to argue that quantum mechanics is not a complete physical theory. It is sometimes referred to as the EPRB paradox for David Bohm, who converted the original thought experiment into something closer to being experimentally testable.

The EPR experiment yields the following dichotomy:

(i) Either the result of a measurement performed on one part A of a quantum system has a non-local effect on the physical reality of another distant part B, in the sense that quantum mechanics can predict outcomes of some measurements carried out at B, or

(ii) Quantum mechanics is incomplete in the sense that some element of physical reality corresponding to B cannot be accounted for by quantum mechanics (that is, some extra variable is needed to account for it.)

Although originally devised as a thought experiment that would demonstrate the incompleteness of quantum mechanics, actual experimental results ironically refutes the principle of locality, invalidating the EPR trio’s

Incompleteness can be understood in two fundamentally different ways:

(i) QM is incomplete because it is not the ‘right’ theory; the right theory would provide descriptive categories to account for all observable behavior and not leave ‘anything to chance’.

(ii) QM is incomplete, but it accurately reflects the way nature is.

Incompleteness understood as (i) is now considered highly controversial, since it contradicts the impossibility of a hidden variables theory which is shown by Bell test experiments. There are many variants of (ii) which is widely considered to be the more orthodox view of quantum mechanics.

Local realism is the combination of the principle of locality with the ‘realistic’ assumption that all objects must objectively have their properties already before these properties are observed. Einstein liked to say that the Moon is ‘out there’ even when no one is observing it.

Local realism is a significant feature of classical mechanics, general relativity and
original purpose. The “spooky action at a distance” that so disturbed the authors of EPR consistently occurs in numerous and widely replicated experiments. Einstein never really accepted quantum mechanics as a ‘real’

Maxwell’s theory, but quantum mechanics largely rejects this principle due the presence of distant quantum entanglements, most clearly demonstrated by the EPR paradox and quantified by Bell’s inequalities. Every theory that, like quantum mechanics, is compatible with violations of Bell’s inequalities must abandon either local realism or counterfactual definiteness. (The vast majority of physicists believe that experiments have demonstrated Bell’s violations, but some local realists dispute the claim, in view of the recognized loopholes in the tests.) Different interpretations of quantum mechanics reject different parts of local realism and/or counterfactual definiteness.

In most of the conventional interpretations, such as the version of the Copenhagen interpretation and the interpretation based on Consistent Histories, where the wave function is not assumed to have a direct physical interpretation or reality it is realism that is rejected. The actual definite properties of a physical system ‘do not exist’ prior to the measurement and the wave function has a restricted interpretation as nothing more than a mathematical tool used to calculate the probabilities of experimental outcomes, in agreement with positivism in philosophy as the only topic that science should discuss.

In the version of the Copenhagen interpretation where the wave function is assumed to have a physical interpretation or reality (the nature of which is unspecified), the principle of locality is violated during the measurement process via wave function collapse. This is a nonlocal process because Born’s Rule, when applied to the system’s wave function, yields a probability density for all regions of space and time. Upon measurement of the physical system, the probability density vanishes everywhere instantaneously, except where (and when) the measured entity is found to exist. This “vanishing” would be a real physical process, and clearly non-local (faster-than-light-speed), if the wave function is considered physically real and the probability density converged to zero at arbitrarily far distances during the finite time required for the measurement process.

The Bohm interpretation always wants to preserve realism, and it needs to violate the principle of locality to achieve the required correlations.

In the many-worlds interpretation realism and locality are retained but counterfactual definiteness is rejected by the extension of the notion of reality to allow the existence of parallel universes.

Because the differences between the different interpretations are mostly philosophical ones (except for the Bohm and many-worlds interpretations), the physicists usually use the language in which the important statements are independent of the interpretation we choose. In this framework, only the measurable action at a distance, a super-luminal propagation of real, physical information, would be usually considered to be a violation of locality by the physicists. Such phenomena have never been seen, and they are not predicted by the current theories (with the possible exception of the Bohm theory).

Locality is one of the axioms of relativistic quantum field theory, as required for causality. The formalization of locality in this case is as follows: if we have two observables, each localized within two distinct space-time regions which happen to be at a space-like separation from each other, the observables must commute. This interpretation of the word ‘locality’ is closely related to the relativistic version in physics. In physics a solution is local if the underlying equations are either Lorentz invariant or, more generally, generally covariant or locally Lorentz invariant.
and complete theory, struggling to the end of his career (and life) for an interpretation that could comply with his Relativity without implying “God playing dice,” as he condensed his dissatisfaction with QM’s intrinsic randomness and (still to be resolved) counter–intuitivity.

The EPR paradox is a paradox in the following sense: if one takes quantum mechanics and adds some seemingly reasonable conditions (referred to as locality, realism, counter factual definiteness, and completeness), then one obtains a contradiction. However, quantum mechanics by itself does not appear to be internally inconsistent, nor does it contradict relativity. As a result of further theoretical and experimental developments since the original EPR paper, most physicists today regard the EPR paradox as an illustration of how quantum mechanics violates classical intuitions, and not as an indication that quantum mechanics is fundamentally flawed.

The EPR paradox draws on a quantum entanglement phenomenon, to show that measurements performed on spatially separated parts of a quantum system can apparently have an instantaneous influence on one another. This effect is now known as nonlocal behavior. In order to illustrate this, let us consider a simplified version of the EPR thought experiment due to David Bohm.

Measurements on an Entangled State

We have a source that emits pairs of electrons, with one electron sent to destination A, where there is an observer named Alice, and another is sent to destination B, where there is an observer named Bob. According

A physical theory is said to ‘exhibit nonlocality’ if, in that theory, it is not possible to treat widely separated systems as independent. The term is most often reserved, however, for interactions that occur outside the backward light cone, i.e. super–luminal influences. Nonlocality does not imply a lack of causality only in the case when ‘ethereal’, not ‘causal’, information is transmitted between systems.

Special Relativity shows that in the case where causal information is transmitted at super–luminal rates, causality is violated. For example, if information could be exchanged at super–luminal rates, it would be possible to arrange for your grandfather to be killed before you are born, which leads to causal paradoxes. Some effects that appear nonlocal in quantum mechanics may actually obey locality, such as quantum entanglement. These interactions effect correlations between states of particles (expressed by a wave ψ–function which may be in a superposition of states), such as the infamous singlet state. Einstein criticized this interpretation of quantum mechanics on the grounds that these effects employed what he called “spooky action at a distance.”

This issue is very closely related to Bell’s theorem and the EPR paradox. Quantum field theory, on the other hand, which is the relativistic generalization of quantum mechanics, contains mathematical features that assure locality.
Fig. 7.2 The EPR thought experiment, performed with electrons. A source (center) sends electrons toward two observers, Alice (left) and Bob (right), who can perform spin measurements.

to quantum mechanics, we can arrange our source so that each emitted electron pair occupies a quantum state called a spin singlet. This can be viewed as a quantum superposition of two states, which we call I and II. In state I, electron A has spin pointing upward along the $z$–axis ($+z$) and electron B has spin pointing downward along the $z$–axis ($−z$). In state II, electron A has spin $−z$ and electron B has spin $+z$. Therefore, it is impossible to associate either electron in the spin singlet with a state of definite spin. The electrons are thus said to be entangled.

Alice now measures the spin along the $z$–axis. She can get one of two possible outcomes: $+z$ or $−z$. Suppose she gets $+z$. According to quantum mechanics, the quantum state of the system collapses into state I (different interpretations of quantum mechanics have different ways of saying this, but the basic result is the same). The quantum state determines the probable outcomes of any measurement performed on the system. In this case, if Bob subsequently measures spin along the $z$–axis, he will get $−z$ with 100% probability. Similarly, if Alice gets $−z$, Bob will get $+z$.

There is nothing special about our choice of the $z$ axis. For instance, suppose that Alice and Bob now decide to measure spin along the $x$–axis. According to quantum mechanics, the spin singlet state may equally well be expressed as a superposition of spin states pointing in the $x$–direction. We will call these states Ia and IIa. In state Ia, Alice’s electron has spin $+x$ and Bob’s electron has spin $−x$. In state IIa, Alice’s electron has spin $−x$ and Bob’s electron has spin $+x$. Therefore, if Alice measures $+x$, the system collapses into Ia, and Bob will get $−x$. If Alice measures $−x$, the system collapses into IIa, and Bob will get $+x$.

In quantum mechanics, the $x$–spin and $z$–spin are incompatible ob-
servables, which means that there is a Heisenberg uncertainty principle operating between them: a quantum state cannot possess a definite value for both variables. Suppose Alice measures the $z$-spin and obtains $+z$, so that the quantum state collapses into state I. Now, instead of measuring the $z$-spin as well, Bob measures the $x$-spin. According to quantum mechanics, when the system is in state I, Bob’s $x$-spin measurement will have a 50% probability of producing $+x$ and a 50% probability of $-x$. Furthermore, it is fundamentally impossible to predict which outcome will appear until Bob actually performs the measurement.

So how does Bob’s electron know, at the same time, which way to point if Alice decides (based on information unavailable to Bob) to measure $x$ and also how to point if Alice measures $z$? Using the usual Copenhagen interpretation rules that say the wave function collapses at the time of measurement, there must be action at a distance or the electron must know more than it is supposed to. To make the mixed part quantum and part classical descriptions of this experiment local, we have to say that the notebooks (and experimenters) are entangled and have linear combinations of $+$ and $-$ written in them, like Schrödinger’s Cat.

Incidentally, although we have used spin as an example, many types of physical quantities — what quantum mechanics refers to as quantum observables, can be used to produce quantum entanglement. The original EPR paper used momentum for the observable. Experimental realizations of the EPR scenario often use the polarization of photons, because polarized photons are easy to prepare and measure.

In recent years, doubt has been cast on EPR’s conclusion due to developments in understanding locality and especially quantum decoherence. The word locality has several different meanings in physics. For example, in quantum field theory locality means that quantum fields at different points of space do not interact with one another. However, quantum field theories that are local in this sense appear to violate the principle of locality as defined by EPR, but they nevertheless do not violate locality in a more general sense. A wave ψ-function collapse can be viewed as an epiphe-
nomenon of quantum decoherence\textsuperscript{27} which in turn is nothing more than an effect of the underlying local time evolution of the wavefunction of a system and all of its environment. Since the underlying behaviour doesn’t violate local causality it follows that neither does the additional effect of wavefunction collapse, whether real or apparent. Therefore, as outlined in the example above, the EPR experiment (nor any quantum experiment) does not demonstrate that FTL signalling is possible.

accommodated into the mathematical formulation of quantum mechanics by postulating that there were two processes of wave $\psi$–function change:

1. The probabilistic, non–unitary, non–local, discontinuous change brought about by observation and measurement, as outlined above.

2. The deterministic, unitary, continuous time evolution of an isolated system that obeys Schrödinger’s equation (or nowadays some relativistic, local equivalent).

In general, quantum systems exist in superpositions of those basis states that most closely correspond to classical descriptions, and – when not being measured or observed, evolve according to the time dependent Schrödinger equation, relativistic quantum field theory or some form of quantum gravity or string theory, which is process (2) mentioned above. However, when the wavefunction collapses – process (1) – from an observer’s perspective the state seems to ‘leap’ or ‘jump’ to just one of the basis states and uniquely acquire the value of the property being measured, e.i., that is associated with that particular basis state. After the collapse, the system begins to evolve again according to the Schrödinger equation or some equivalent wave equation.

Hence, in experiments such as the double–slit experiment each individual photon arrives at a discrete point on the screen, but as more and more photons are accumulated, they form an interference pattern overall.

Consciousness causes collapse is the theory that observation by a conscious observer is responsible for the wave $\psi$–function collapse. It is an attempt to solve the Wigner’s friend paradox by simply stating that collapse occurs at the first ‘conscious’ observer. Supporters claim this is not a revival of substance dualism, since (in a ramification of this view) consciousness and objects are entangled and cannot be considered as separate. The consciousness causes collapse theory can be considered as a speculative appendage to almost any interpretation of quantum mechanics and many physicists reject it as unverifiable and introducing unnecessary elements into physics.

In recent decades the latter view has gained popularity.

\textsuperscript{27}Quantum decoherence is the mechanism by which quantum systems interact with their environments to exhibit probabilistically additive behaviour (a feature of classical physics) and give the appearance of wavefunction collapse. Decoherence occurs when a system interacts with its environment, or any complex external system, in such a thermodynamically irreversible way that ensures different elements in the quantum superposition of the system + environment’s wave $\psi$–function can no longer interfere with each other.

Decoherence does not provide a mechanism for the actual wave function collapse; the quantum nature of the system is simply ‘leaked’ into the environment so that a total superposition of the wavefunction still exists, but exists beyond the realm of measurement; rather decoherence provides a mechanism for the appearance of wavefunction collapse.

Decoherence represents a major problem for the practical realization of quantum computers, since these heavily rely on the undisturbed evolution of quantum coherences.
Resolving the Paradox

There are several ways to resolve the EPR paradox. The one suggested by EPR is that quantum mechanics, despite its success in a wide variety of experimental scenarios, is actually an incomplete theory. In other words, there is some yet undiscovered theory of nature to which quantum mechanics acts as a kind of statistical approximation (albeit an exceedingly successful one). Unlike quantum mechanics, the more complete theory contains variables corresponding to all the ‘elements of reality’. There must be some unknown mechanism acting on these variables to give rise to the observed effects of ‘non–commuting quantum observables’, i.e., the Heisenberg uncertainty principle. Such a theory is called a hidden variable theory.

Another resolution of the EPR paradox is provided by Bell’s theorem.

7.1.2.2 Bell’s Theorem

Bell’s theorem is the most famous legacy of the late physicist John Bell. It is notable for showing that the predictions of quantum mechanics (QM) differ from those of intuition. It is simple and elegant, and touches upon fundamental philosophical issues that relate to modern physics. In its simplest form, Bell’s theorem states:

No physical theory of local hidden variables can ever reproduce all of the predictions of quantum mechanics.

This theorem has even been called “the most profound in science” [Stapp (1995)]. Bell’s seminal 1964 paper was entitled “On the Einstein Podolsky Rosen paradox”. The Einstein Podolsky Rosen paradox (EPR paradox) assumes local realism, the intuitive notion that particle attributes have definite values independent of the act of observation and that physical effects have a finite propagation speed. Bell showed that local realism leads...
to a requirement for certain types of phenomena that are not present in quantum mechanics. This requirement is called Bell’s inequality.

Different authors subsequently derived similar inequalities, collectively termed Bell inequalities, that also assume local realism. That is, they assume that each quantum-level object has a well defined state that accounts for all its measurable properties and that distant objects do not exchange information faster than the speed of light. These well defined properties are often called hidden variables, the properties that Einstein posited when he stated his famous objection to quantum mechanics: “God does not play dice.”

The inequalities concern measurements made by observers (often called Alice and Bob) on entangled pairs of particles that have interacted and then separated. Hidden variable assumptions limit the correlation of subsequent measurements of the particles. Bell discovered that under quantum mechanics this correlation limit may be violated. Quantum mechanics lacks local hidden variables associated with individual particles, and so the inequalities do not apply to it. Instead, it predicts correlation due to quantum entanglement of the particles, allowing their state to be well defined only after a measurement is made on either particle. That restriction agrees with the Heisenberg uncertainty principle, one of the most fundamental concepts in quantum mechanics.

Per Bell’s theorem, either quantum mechanics or local realism is wrong. Experiments were needed to determine which is correct, but it took many years and many improvements in technology to perform them.

Bell test experiments to date overwhelmingly show that the inequalities of Bell’s theorem are violated. This provides empirical evidence against local realism and demonstrates that some of the “spooky action at a distance” suggested by the famous Einstein Podolsky Rosen (EPR) thought experiment do in fact occur. They are also taken as positive evidence in favor of QM. The principle of special relativity is saved by the no-communication theorem, which proves that the observers cannot use the inequality violations to communicate information to each other faster than the speed of light.

John Bell’s papers examined both John von Neumann’s 1932 proof of the incompatibility of hidden variables with QM and Albert Einstein and his colleagues’ seminal 1935 paper on the subject.
Importance of the Theorem

After EPR, quantum mechanics was left in the unsatisfactory position that it was either incomplete in the sense that it failed to account for some elements of physical reality, or it violated the principle of finite propagation speed of physical effects. In the EPR thought experiment, two observers, now commonly referred to as Alice and Bob, perform independent measurements of spin on a pair of electrons, prepared at a source in a special state called a spin singlet state. It was a conclusion of EPR that once Alice measured spin in one direction (e.g. on the x axis), Bob’s measurement in that direction was determined with certainty, whereas immediately before Alice’s measurement, Bob’s outcome was only statistically determined. Thus, either the spin in each direction is not an element of physical reality or the effects travel from Alice to Bob instantly.

In QM predictions were formulated in terms of probabilities, for example, the probability that an electron might be detected in a particular region of space, or the probability that it would have spin up or down. However, there still remained the idea that the electron had a definite position and spin, and that QM’s failing was its inability to predict those values precisely. The possibility remained that some yet unknown, but more powerful theory, such as a hidden variable theory, might be able to predict these quantities exactly, while at the same time also being in complete agreement with the probabilistic answers given by QM. If a hidden variables theory were correct, the hidden variables were not described by QM and thus QM would be an incomplete theory.

The desire for a local realist theory was based on two ideas: first, that objects have a definite state that determines the values of all other measurable properties such as position and momentum and second, that (as a result of special relativity) effects of local actions such as measurements cannot travel faster than the speed of light. In the formalization of local realism used by Bell, the predictions of a theory result from the application of classical probability theory to an underlying parameter space. By a simple (but clever) argument based on classical probability he then showed that correlations between measurements are bounded in a way that is violated by QM.

Bell’s theorem seemed to seal the fate of those that had local realist hopes for QM.
Bell’s Thought Experiment

Bell considered a setup in which two observers, Alice and Bob, perform independent measurements on a system S prepared in some fixed state. Each observer has a detector with which to make measurements. On each trial, Alice and Bob can independently choose between various detector settings. Alice can choose a detector setting a to get a measurement A(a) and Bob can choose a detector setting b to measure B(b). After repeated trials Alice and Bob collect statistics on their measurements and correlate the results.

There are two key assumptions in Bell’s analysis: (1) each measurement reveals an objective physical property of the system (2) a measurement taken by one observer has no effect on the measurement taken by the other.

In the language of probability theory, repeated measurements of system properties can be regarded as repeated sampling of random variables. One might expect measurements by Alice and Bob to be somehow correlated with each other: the random variables are assumed not to be independent, but linked in some way. Nonetheless, there is a limit to the amount of correlation one might expect to see. This is what the Bell inequality expresses.

A version of the Bell inequality appropriate for this example is given by John Clauser, Michael Horne, Abner Shimony and R. A. Holt, and is called the CHSH form,

\[ C[A(a), B(b)] + C[A(a), B(b')] + C[A(a'), B(b)] - C[A(a'), B(b')] \leq 2, \]

where \( C \) denotes correlation.

Description of the Bell’s Theorem

Continuing on from the situation explored in the EPR paradox, consider that again a source produces paired particles, one sent to Alice and another to Bob. When Alice and Bob measure the spin of the particles in the same axis, they will get identical results; when Bob measures at right angles to Alice’s measurements they will get the same results 50% of the time, the same as a coin toss. This is expressed mathematically by saying that in the first case, their results have a correlation of 1, or perfect correlation; in the second case they have a correlation of 0; no correlation (a correlation of \(-1\) would indicate getting opposite results the whole time).
So far, this can be explained by positing local hidden variables; each pair of particles is sent out with instructions on how to behave when measured in the $x$–axis and the $z$–axis, generated randomly. Clearly, if the source only sends out particles whose instructions are correlated for each axis, then when Alice and Bob measure on the same axis, they are bound to get identical results; but (if all four possible pairs of instructions are generated equally) when they measure on perpendicular axes they will see zero correlation.

Now consider that B rotates their apparatus (by 45 degrees, say) relative to that of Alice. Rather than calling the axes $x_A$, etc., henceforth we will call Alice’s axes $a$ and $a'$, and Bob’s axes $b$ and $b'$. The hidden variables (supposing they exist) would have to specify a result in advance for every possible direction of measurement. It would not be enough for the particles to decide what values to take just in the direction of the apparatus at the time of leaving the source, because either Alice or Bob could rotate their apparatus by a random amount any time after the particles left the source.

Next, we define a way to ‘keep score’ in the experiment. Alice and Bob decide that they will record the directions they measured the particles in, and the results they got; at the end, they will tally up, scoring +1 for each time they got the same result and −1 for an opposite result, except that if Alice measured in $a$ and Bob measured in $b'$, they will score +1 for an opposite result and −1 for the same result. It turns out (see the mathematics below) that however the hidden variables are contrived, Alice and Bob cannot average more than 50% overall. For example, suppose that for a particular value of the hidden variables, the $a$ and $b$ directions
are perfectly correlated, as are the $a'$ and $b'$ directions. Then, since $a$ and $a'$ are at right angles and so have zero correlation, $a'$ and $b$ have zero correlation, as do $a$ and $b'$. The unusual ‘scoring system’ is designed in part to ensure this holds for all possible values of the hidden variables.

The question is now whether Alice and Bob can score higher if the particles behave as predicted by quantum mechanics. It turns out they can; if the apparatuses are rotated at 45° to each other, then the predicted score is 71%. In detail: when observations at an angle of $\theta$ are made on two entangled particles, the predicted correlation between the measurements is $\cos \theta$. In one explanation, the particles behave as if when Alice makes a measurement (in direction $x$, say), Bob’s particle instantaneously switches to take that direction. When Bob makes a measurement, the correlation (the averaged–out value, taking +1 for the same measurement and −1 for the opposite) is equal to the length of the projection of the particle’s vector onto his measurement vector; by trigonometry, $\cos \theta$. $\theta$ is 45°, and $\cos \theta$ is $\sqrt{2}/2$, for all pairs of axes except $(a, b')$, where they are 135° and $-\sqrt{2}/2$, but this last is taken in negative in the agreed scoring system, so the overall score is $\sqrt{2}/2$; 0.707, or 71%. If experiment shows (as it appears to) that the 71% score is attained, then hidden variable theories cannot be correct; not unless information is being transmitted between the particles faster than light, or the experimental design is flawed.

For the mathematical statement and proof of the Bell’s theorem, see [Bell (1964)]; [Bell (1966)]; [Bell (1987)].

7.1.3 Orchestrated Reduction and Penrose Paradox

Orchestrated Objective Reduction (Orch OR) is a theory of consciousness put forth in the mid–1990s by British mathematical physicist Sir Roger Penrose and American anesthesiologist Stuart Hameroff (see [Hameroff and Watt (1983)]; [Hameroff and Watt (1982)]). Whereas some theories assume consciousness emerges from the brain, and among these some assume that mind emerges from complex computation at the level of synapses among brain neurons, Orch OR involves a specific form of quantum computation which underlies these neuronal synaptic activities. The proposed quantum computations occur in structures inside the brain’s neurons called microtubules.

Now, recall from above that to make a measurement or observation of the quantum system means to concentrate on those aspects of the system that can be simultaneously magnified to the classical level and from which
the system must then choose. Therefore, by measuring we are disturbing the quantum system with the magnifying device, which results in decoherence. In other words, we get classical probabilities, highly reminiscent of a standard particle-like behavior. The ‘measurement/observation’ process has caused decoherence of the wave $\psi$–function and thus led to its collapse to a specific state.

Until now, our approach to the quantum world involves two components: the one component dubbed by Penrose the $U$–part, involves the unitary evolution of the system, in a deterministic, continuous, time–symmetric fashion, as described for example by the Schrödinger equation (2.3), i.e.,

$$ U : \quad i\hbar \frac{\partial}{\partial t} \psi(t) = \hat{H} \psi(t). \quad (7.1) $$

Clearly such an evolution respects quantum coherence, as reflected by the quantum complex superposition principle implicit in (7.1). The second component, dubbed by Penrose the $R$–part, involves the reduction of the state–vector or collapse of the wave $\psi$–function, that enforces coexisting alternatives to resolve themselves into actual alternatives, one or the other,

$$ R : \quad \psi = \sum_i \epsilon_i \psi_i \longrightarrow \sum_i |c_i|^2 \psi_i^2, \quad (7.2) $$

where $|c_i|^2$ are classical probabilities describing actual alternatives. It is the $R$–part of quantum physics that introduces ‘uncertainties’ and ‘probabilities’, thus involving discontinuous, time–asymmetric quantum jumps and leading to gross violations of quantum coherence. It is fair to say that almost universally, when physicists talk about quantum physics, they tacitly identify it with its $U$–part only. It is the $U$–part that has absorbed all our attention for about 70 years now, and in its more advanced form, relativistic quantum field theory, has become an icon of modern physics, with spectacular success, e.g., the Standard Model $SU(3) \times SU(2) \times U(1)$. On the other hand, the $R$–part has been vastly and conveniently forgotten, tacitly assumed to be some mere technicality that gives us the right rules of ‘measurement’ or ‘observation’: different aspects of a quantum system are simultaneously magnified at the classical level, and between which the system must choose. This attitude has brought us finally to the Penrose paradox, and we need to reconsider our strategy. Actually, there is no way to deduce the $R$–part from the $U$–part, the $R$–part being a completely different procedure from the $U$–part, and effectively providing the other ‘half’ of the interpretation of quantum mechanics. It is the $(U + R)$ parts together that
are needed for the spectacular agreement of quantum mechanics with the observed facts. So, one is after some new dynamics, $N$, that provides a unified and comprehensible picture of the whole $(U+R)$ process. In the work of [Ellis et al. (1992); Ellis et al. (1999); Mavromatos and Nanopoulos (1995a); Mavromatos and Nanopoulos (1995b); Nanopoulos (1995)] the above approach is presented by

$$U \oplus R \subseteq N.$$ (7.3)

It should be stressed that the new dynamics involved in the $N$–equation (7.3), because they have to approach at appropriate limits the $U$–equation (7.1) and the $R$–equation (7.2), i.e., almost anti–diametrical points of view, cannot be some smooth generalization of some wave dynamics. Apparently, the $N$–equation (7.3) has to contain seeds of non–relativistic invariance and time asymmetry, but in such a way that when the $R$–part or emergence of classicality is neglected, an approximately relativistic, time–symmetric (quantum field) theory emerges.

7.1.3.1 Orchestrated Objective Reduction

In the 1970s and 1980s Hameroff attempted to show that consciousness depends on computation within neurons in microtubules, self–assembling cylindrical polymers of the protein tubulin. Microtubules organize neuronal shape and function, e.g., forming and maintaining synapses (and help single cells like paramecium swim, find food and mates, learn and have sex without any synapses). Hameroff concluded that microtubules function as molecular–level cellular automata, and that microtubules in each neuron of the brain had the computational power of $10^{16}$ operations per second. Neuronal–level synaptic operations were regulated by these internal computations, Hameroff claimed, so attempts by artificial intelligence (AI) workers to mimic brain functions by simulating neuronal/synaptic activities would fail. Hence, as far as explaining consciousness, why we have inner experience, feelings, subjectivity, merely adding another layer of information processing within neurons in microtubules did not help.

Meanwhile Roger Penrose, famous for his work in relativity, quantum mechanics, geometry and other disciplines, had concluded for completely different reasons that AI computational approaches were inadequate to explain consciousness. In his 1989 book ‘The Emperor’s New Mind’ Penrose used Kurt Gödel’s theorem to argue that human consciousness and understanding required a factor outside algorithmic computation, and that the
missing ‘non–computable’ factor was related to a specific type of quantum computation involving what he termed ‘objective reduction’ (OR), his solution to the measurement problem in quantum mechanics.

Penrose considered superposition as a separation in underlying reality at its most basic level, the Planck scale. Tying quantum superposition to general relativity, he identified superposition as space–time curvatures in opposite directions, hence a separation in fundamental space–time geometry. However, according to Penrose, such separations are unstable and will reduce at an objective threshold, hence avoiding multiple universes.

The threshold for Penrose OR is given by the indeterminacy principle \( E = \hbar / t \), where \( E \) is the gravitational self–energy (i.e., the degree of space–time separation given by the superpositioned mass), \( \hbar \) is Planck’s constant over \( 2\pi \), and \( t \) is the time until OR occurs. Thus the larger the superposition, the faster it will undergo OR, and vice versa. Small superpositions, e.g., an electron separated from itself, if isolated from environment would require 10 million years to reach OR threshold. An isolated one kilogram object (e.g., Schrodinger’s cat) would reach OR threshold in only 10–37 seconds. Penrose OR is currently being tested.

An essential feature of Penrose OR is that the choice of states when OR occurs is selected neither randomly (as are choices following measurement, or decoherence) nor completely algorithmically. Rather, states are selected by a ‘non–computable’ influence involving information embedded in the fundamental level of space–time geometry at the Planck scale. Moreover, Penrose claimed that such information is Platonic, representing pure mathematical truth, aesthetic and ethical values. Plato had proposed such pure values and forms, but in an abstract realm. Penrose placed the Platonic realm in the Planck scale.

In ‘The Emperor’s New Mind’ [Penrose (1967)], Penrose suggested (and further developed later in [Penrose (1994)] [Penrose (1997)]) that consciousness required a form of quantum computation in the brain.

Recall that quantum computation had been suggested by Richard Feynman, Paul Benioff and David Deutsch in the 1980s. The idea is that classical information (e.g., bit states of either 1 or 0) could also be quantum superpositions of both 1 and 0 (quantum bits, or qubits). Such qubits interact and compute by nonlocal quantum entanglement, eventually being measured/observed and reducing to definite states as the solution. Quantum computations were shown to have enormous capacity if they could be constructed e.g., using qubits of ion states, electron spin, photon polarization, current in Josephson junction, quantum dots, etc. During quantum com-
putation, qubits must be isolated from environmental interaction to avoid loss of superposition, i.e., ‘decoherence’.

Penrose argued that quantum computation which terminated not by measurement, but by his version of objective reduction, constituted consciousness (allowing Platonic non-computable influences). Penrose had no definite biological qubits for such quantum computation by OR, except to suggest the possibility of superpositions of neurons both ‘firing and not firing’.

Hameroff read ‘The Emperor’s New Mind’ and suggested to Penrose that microtubules within neurons were better suited for quantum computing with OR than were superpositions of neuronal firings. The two met in the early 1990s and began to develop the theory now known as Orch OR. ‘Orch’ refers to orchestration, the manner in which biological conditions including synaptic–level neuronal events provide feedback to influence quantum computation with OR in microtubules [Hameroff (1987); Hameroff and Penrose (1996); Hameroff and Penrose (1993); Hameroff (1998)].

7.1.3.2 The Orch OR Model

For biological qubits, Penrose and Hameroff chose conformational states of the tubulin subunit proteins in microtubules. Tubulin qubits would interact and compute by entanglement with other tubulin qubits in microtubules in the same and different neurons.

It was known that the peanut–shaped tubulin protein flexes 30 degrees, giving two different conformational shapes. Could such different states exist as superpositions, and if so, how? Penrose and Hameroff considered three possible types of tubulin superpositions: separation at the level of the entire protein, separation at the level of the atomic nuclei of the individual atoms within the proteins, and separation at the level of the protons and neutrons (nucleons) within the protein. Calculating the gravitational self–energy $E$ of the three types, separation at the level of atomic nuclei had the highest energy, and would be the dominant factor. Penrose and Hameroff calculated $E$ for superposition/separation of one tubulin qubit at the level of atomic nuclei in all the amino acids of the protein. They then related this to brain electrophysiology.

There are claims that the best electrophysiological correlate of consciousness can be seen in the so–called gamma–EEG waves, synchronized oscillations in the range of 30 to 90 Hz (also known as ‘coherent 40 Hz’).
mediated by dendritic membrane depolarizations (not axonal action potentials). This means that roughly 40 times per second (every 25 msec) neuronal dendrites depolarize synchronously throughout wide regions of brain. On the other hand, there are also claims that relaxed alpha waves (8–12 Hz) or even meditative theta waves (4–8 Hz), which have low frequency but high amplitude – are the real source of human creativity.

Using the indeterminacy principle \( E = \hbar / t \) for OR, Penrose and Hameroff used 25 msec for \( t \), and calculated \( E \) in terms of number of tubulins (since \( E \) was known for one tubulin). Thus they were asking: how many tubulins would be required to be in isolated superposition to reach OR threshold in 25 msec, 40 times per second, corresponding with membrane–level brain–wide effects? The answer turned out to be \( 2 \times 10^{11} \) tubulins.

There are roughly 107 tubulins per neuron. If all tubulins in microtubules in a given neuron were involved, this would correspond with \( 2 \times 10^4 \) (20,000) neurons. However, because dendrites are apparently more involved in consciousness than axons (which contain many microtubules), and because not all microtubules in a given dendrite are likely to be involved at any one time, an estimate of, say, 10 percent involvement gives 200,000 neurons involved in consciousness every 25 msec. These estimates (20,000 to 200,000 neurons) fit very well with others from more conventional approaches suggesting tens to hundreds of thousands of neurons are involved in consciousness at any one time.

How would microtubule quantum superpositions avoid environmental decoherence? Cell interiors are known to alternate between liquid phases (solution: ‘sol’) and quasi–solid (gelatinous: ‘gel’) phases due to polymerization states of the ubiquitous protein actin. In the actin-polymerized gel phase, cell water and ions are ordered on actin surfaces, so microtubules are embedded in a highly structured (i.e., non–random) medium. Tubulins are also known to have C termini ‘tails’, negatively charged peptide sequences extending string–like from the tubulin body into the cytoplasm, attracting positive ions and forming a plasma–like Debye layer which can also shield microtubule quantum states. Finally, tubulins in microtubules were suggested to be coherently pumped laser–like into quantum states by biochemical energy (as proposed by Herbert Frohlich).

Actin gelation cycling with 40 Hz events permits input to, and output from isolated microtubule quantum states. Thus during classical, liquid phases of actin depolymerization, inputs from membrane/synaptic inputs could ‘orchestrate’ microtubule states. When actin gelation occurs, quan-
tum isolation and computation ensues until OR threshold is reached, and actin depolymerizes. The result of each OR event (in terms of patterns of tubulin states) would proceed to organize intraneuronal activities including axonal firing and synaptic modulation/learning. Each OR event (e.g., 40 per second) is proposed to be a conscious event, equivalent in philosophical terms to what philosopher Alfred North Whitehead called ‘occasions of experience’.

Thus one implication of the Orch OR model is that consciousness is a sequence of discrete events, rather than a continuum. Yet conscious experience is subjectively uninterrupted, analogous to a movie appearing continuous to observers despite being a series of frames. The difference is that in Orch OR, each conscious event is itself an intrinsic, subjective observation. Moreover the frequency of conscious events may vary, 40 Hz being an average. If someone is excited and conscious events occur more often, (e.g., at 60 Hz), then subjectively the external world seems slower, as great athletes report during peak performance. By $E = \hbar/t$, more frequent conscious events correspond with greater $E$, hence more tubulins/neurons per conscious events and greater intensity of experience. Thus a spectrum of conscious events may exist, similar to photons. There exists a spectrum of conscious quanta–like events ranging from longer wavelength, low intensity events (large $t$, low $E$) and shorter wavelength, higher intensity events (small $t$, large $E$).

7.1.4 Physical Science and Consciousness

Among all the human endeavors, physical science is usually considered to be the most powerful for the maximum power it endows us to manipulate the nature through an understanding of our position in it. This understanding is gained when a set of careful observations based on tangible perceptions, acquired by sensory organs and/or their extensions, is submitted to the logical analysis of human intellect as well as to the intuitive power of imagination to yield the abstract fundamental laws of nature that are not self–evident at the gross level of phenomenal existence. There exists a unity in nature at the level of laws that corresponds to the manifest diversity at the level of phenomena [Samal (1999)].

Can consciousness be understood in this sense by an appropriate use of the methodology of science? The most difficult problem related to consciousness is perhaps, ‘how to define it?’ Consciousness has remained a unitary subjective experience, its various ‘components’ being reflective (the
recognition by the thinking subject of its own actions and mental states), perceptual (the state or faculty of being mentally aware of external environment) and a free will (volition). But how these components are integrated to provide the unique experience called ‘consciousness’, familiar to all of us, remains a mystery. Does it lie at the level of ‘perceptions’ or at the level of ‘laws’? Can it be reduced to some basic ‘substance’ or ‘phenomenon’? Can it be manipulated in a controlled way? Is there a need for a change of either the methodology or the paradigm of science to answer the above questions?

7.1.4.1 Reducing the Consciousness to Its Elements

Now, most of the successes of science over the past five hundred years or so can be attributed to the great emphasis it lays on the ‘reductionist paradigm’. Following this approach, can consciousness be reduced either to ‘substance’ or ‘phenomena’ in the sense that by understanding which one can understand consciousness?

**Physical Substratum**

The attempts to reduce consciousness to a physical basis have been made in the following ways by trying to understand the mechanism and functioning of the human brain in various different contexts [Samal (1999)].

- **Physics**

  The basic substratum of physical reality is the ‘state’ of the system and the whole job of physics can be put into a single question: *Given the initial state, how to predict its evolution at a later time?* In classical world, the state and its evolution can be reduced to events and their spatio–temporal correlations. Consciousness has no direct role to play in this process of reduction, although it is responsible to find an ‘objective meaning’ in such a reduction.

  But the situation is quite different in the quantum world as all relevant physical information about a system is contained in its wave $\psi$–function (or equivalently in its state vector), which is not physical in the sense of being directly measurable. Consciousness plays no role in the *deterministic and unitary Schrödinger evolution* (i.e., the $U$–process of Penrose [Penrose (1967)]) that the ‘un–physical’ wave $\psi$–function undergoes.
To extract any physical information from the wave $\psi$–function one has to use the Born–Dirac rule and thus probability enters in a new way into the quantum mechanical description despite the strictly deterministic nature of evolution of the wave $\psi$–function. The measurement process forces the system to choose an ‘actuality’ from all ‘possibilities’ and thus leads to a non–unitary collapse of the general wave $\psi$–function to an eigenstate (i.e., the $\mathbf{R}$–process of Penrose [Penrose (1967)]) of the concerned observable. The dynamics of this $\mathbf{R}$–process is not known and it is here some authors like Wigner have brought in the consciousness of the observer to cause the collapse of the wave $\psi$–function. But instead of explaining the consciousness, this approach uses consciousness for the sake of Quantum Mechanics which needs the $\mathbf{R}$–process along with the $\mathbf{U}$–process to yield all its spectacular successes.

The $\mathbf{R}$–process is necessarily nonlocal and is governed by an irreducible element of chance, which means that the theory is not naturalistic: the dynamics is controlled in part by something that is not a part of the physical universe. Stapp [Stapp (1995)] has given a quantum–mechanical model of the brain dynamics in which this quantum selection process is a causal process governed not by pure chance but rather by a mathematically specified nonlocal physical process identifiable as the conscious process. It was reported that attempts have been made to explain consciousness by relating it to the ‘quantum events’, but any such attempt is bound to be futile as the concept of ‘quantum event’ in itself is ill–defined.

Keeping in view the fundamental role that the quantum vacuum plays in formulating the quantum field theories of all four known basic interactions of nature spreading over a period from the Big–Bang to the present, it has been suggested that if at all consciousness be reduced to anything ‘fundamental’ that should be the ‘quantum vacuum’ in itself. But in such an approach the following questions arise:

1) If consciousness has its origin in the quantum vacuum that gives rise to all fundamental particles as well as the force fields, then why is it that only living things possess consciousness?

2) What is the relation between the quantum vacuum that gives rise to consciousness and the space–time continuum that confines all our perceptions through which consciousness manifests itself?

3) Should one attribute consciousness only to systems consisting of ‘real’ particles or also to systems containing ‘virtual’ particles? Despite these questions, the idea of tracing the origin of ‘consciousness’ to ‘sub-
stantial nothingness’ appears quite promising because the properties of ‘quantum vacuum’ may ultimately lead us to an understanding of the dynamics of the \( R \)-process and thus to a physical comprehension of consciousness.

One of the properties that distinguishes living systems from the non-living systems is their ability of self-organization and complexity. Since life is a necessary condition for possessing consciousness, can one attribute consciousness to a ‘degree of complexity’ in the sense that various degrees of consciousness can be caused by different levels of complexity? Can one give a suitable quantitative definition of consciousness in terms of ‘entropy’ that describes the ‘degree of self-organization or complexity’ of a system? What is the role of non-linearity and non-equilibrium thermodynamics in such a definition of consciousness? In this holistic view of consciousness what is the role played by the phenomenon of quantum nonlocality, first envisaged in EPR paper and subsequently confirmed experimentally [Aspect et al. (1982)]? What is the role of irreversibility and dissipation in this holistic view?

- **Neurobiology**

On the basis of the vast amount of information available on the structure and the modes of communication (neurotransmitters, neuromodulators, neurohormones) of the neuron, neuroscience has empirically found [Samal (1999)] the neural basis of several attributes of consciousness. With the help of modern scanning techniques and by direct manipulations of the brain, neurobiologists have found out that various human activities (both physical and mental) and perceptions can be mapped into almost unique regions of the brain. Awareness, being intrinsic to neural activity, arises in higher level processing centers and requires integration of activity over time at the neuronal level. But there exists no particular region that can be attributed to have given rise to consciousness. Consciousness appears to be a collective phenomena where the ‘whole’ is much more than the sum of parts. Is each neuron having the ‘whole of consciousness’ within it, although it does work towards a particular attribute of consciousness at a time?

Can this paradigm of finding neural correlates of the attributes of consciousness be fruitful in demystifying consciousness? Certainly not. As it was aptly concluded [Samal (1999)] the currently prevalent reductionist approaches are unlikely to reveal the basis of such holistic phenomenon as consciousness. There have been holistic attempts [Hameroff (1987); Penrose (1967)] to understand consciousness
in terms of collective quantum effects arising in cytoskeletons and microtubules; minute substructures lying deep within the brain’s neurons. The effect of general anaesthetics like chloroform (CHCl₃), isofluorane (CHF₂OCHClCF₃) etc. in switching off the consciousness, not only in higher animals such as mammals or birds but also in paramecium, amoeba, or even green slime mould has been advocated [Hameroff and Watt (1983)] to be providing a direct evidence that the phenomenon of consciousness is related to the action of the cytoskeleton and to microtubules. But all the implications of ‘quantum coherence’ regarding consciousness in such approach can only be unfolded after we achieve a better understanding of ‘quantum reality’, which still lies ahead of the present–day physics.

- **AI and CI**

Can machines be intelligent? Within the restricted definition of ‘artificial intelligence’, the neural network approach has been the most promising one. But the possibility of realising a machine capable of artificial intelligence based on this approach is constrained at present [Samal (1999)] by the limitations of ‘silicon technology’ for integrating the desired astronomical number of ‘neuron–equivalents’ into a reasonable compact space. Even though we might achieve such a feat in the foreseeable future by using chemical memories, it is not quite clear whether such artificially intelligent machines can be capable of ‘artificial consciousness’. Because one lacks at present a suitable working definition of ‘consciousness’ within the framework of studies involving artificial intelligence.

**Invoking Gödel’s incompleteness theorem**, Penrose has argued [Penrose (1967)] that the technology of electronic computer–controlled robots will not provide a way to the artificial construction of an actually intelligent machine–in the sense of a machine that ‘understands’ what it is doing and can act upon that understanding. He maintains that human understanding (hence consciousness) lies beyond formal arguments and beyond computability i.e., in the Turing–machine–accessible sense.

Assuming the inherent ability of quantum mechanics to incorporate consciousness, can one expect any improvement in the above situation by considering ‘computation’ to be a physical process that is governed by the rules of quantum mechanics rather than that of classical physics? In ‘Quantum computation’ [Deutsch (1985)] the classical notion of a Turing machine is extended to a corresponding quantum one that takes into account the quantum superposition principle. In ‘standard’ quan-
Quantum Leap

tum computation, the usual rules of quantum theory are adopted, in which the system evolves according to the $U-$process for essentially the entire operation, but the $R-$process becomes relevant mainly only at the end of the operation, when the system is ‘measured’ in order to ascertain either the termination or the result of the computation. Although the superiority of the quantum computation over classical computation in the sense of complexity theory have been shown [Deutsch (1992)], Penrose insists that it is still a ‘computational’ process since $U-$process is a computable operation and $R-$process is purely probabilistic procedure. What can be achieved in principle by a quantum computer could also be achieved, in principle, by a suitable Turing–machine–with–randomizer. Thus he concludes that even a quantum computer would not be able to perform the operations required for human conscious understanding. But we think that such a view is limited because ‘computation’ as a process need not be confined to a Turing–machine–accessible sense and in such situations one has to explore the power of quantum computation in understanding consciousness.

We conclude from the above discussions that the basic physical substrata to which consciousness may be reduced are ‘neuron’, ‘event’ and ‘bit’ at the classical level, whereas at the quantum level they are ‘microtubule’, ‘wave $\psi-$function’ and ‘qubit’; depending on whether the studies are done in neurobiology, physics and computer science respectively [Samal (1999)]. Can there be a common platform for these trio of substrata?

We believe the answer to be in affirmative and the first hint regarding this comes from John Wheeler’s [Wheeler (1989)] remarkable idea: “Every particle, every field of force, even the space-time continuum itself, derives its function, its meaning, its very existence entirely, even if in some contexts indirectly, from the apparatus, elicited answers to yes or no questions, binary choices, bits”. This view of the world refers not to an object, but to a vision of a world derived from pure logic and mathematics in the sense that an immaterial source and explanation lies at the bottom of every item of the physical world. In a report [Wilczek (1999)] the remarkable extent of embodiment of this vision in modern physics has been discussed along with the possible difficulties faced by such a scheme. But can this scheme explain consciousness by reducing it to bits? Perhaps not unless it undergoes some modification.

Because consciousness involves an awareness of an endless mosaic of
The major hurdles to be cleared are:

1. Observer or Participator? In such equipment–evoked, quantum–information–theoretic approach, the inseparability of the observer from the observed will bring in the quantum measurement problem either in the form of dynamics of the $R -$ process or in the emergence of classicality of the world from a quantum substratum. We first need the solutions to these long–standing problems before attempting to reduce the ‘fuzzy’ world of consciousness to ‘qubits’.

2. Communication? Even if we get the solutions to the above problems that enable us to reduce the ‘attributes of consciousness’ to ‘qubits’, still then the ‘dynamics of the process that gives rise to consciousness’ will be beyond ‘quantum information’ as it will require a suitable definition of ‘communication’ in the sense expressed by [Follesdal (1975)]. “Meaning is the joint product of all evidence that is available to those who communicate.” Consciousness helps us to find a ‘meaning’ or ‘understanding’ and will depend upon ‘communication’. Although all ‘evidence’ can be reduced to qubits, ‘communication’ as an exchange of qubits has to be well–defined. Why do we say that a stone or a tree is unconscious? Is it because we do not know how to ‘communicate’ with them? Can one define ‘communication’ in physical terms beyond any verbal or non–verbal language? Where does one look for a suitable definition of ‘communication’? Maybe one has to define ‘communication’ at the ‘substantial nothingness’ level of quantum vacuum.

3. Time’s Arrow? How important is the role of memory in ‘possessing consciousness’? Would our consciousness be altered if the world we experience were reversible with respect to time? Can our consciousness ever find qualitatively different things, such as the color of a rose, the fragrance of a perfume, the music of a piano, the tactile sense of objects, the power of abstraction, the intuitive feeling for time and space, emotional states like love and hate, the ability to put oneself in other’s position, the ability to wonder, the power to wonder at one’s wondering etc. It is almost impossible to reduce them all to the 0–or–1 sharpness of the definition of ‘bits’. A major part of human experience and consciousness is fuzzy and hence can not be reduced to yes or no type situations. Hence we believe that ‘bit’ has to be modified to incorporate this fuzzyness of the world. Perhaps the quantum superposition inherent to a ‘qubit’ can help. Can one then reduce the consciousness to a consistent theory of ‘quantum information’ based on qubits? Quite unlikely, till our knowledge of ‘quantum reality’ and the ‘emergence of classicality from it’ becomes more clear.
out why it is not possible to influence the past?

Hence we conclude that although consciousness may be beyond ‘computability’, it is not beyond ‘quantum communicability’ once a suitable definition for ‘communication’ is found that exploits the quantum superposition principle to incorporate the fuzzyness of our experience. Few questions arise:

(1) How to modify the qubit?

(2) Can a suitable definition of ‘communication’, based on immaterial entity like ‘qubit’ or ‘modified qubit’, take care of non–physical experience like dream or thoughts? We assume, being optimistic, that a suitable modification of ‘qubit’ is possible that will surpass the hurdles of communicability, dynamics of $\mathbb{R}$–process and irreversibility. For the lack of a better word we will henceforth call such a modified qubit as ‘Basic Entity’ (BE) [Samal (1999)].

Non–Physical Substratum

Unlike our sensory perceptions related to physical ‘substance’ and ‘phenomena’ there exists a plethora of human experiences like dreams, thoughts and lack of any experience during sleep which are believed to be non–physical in the sense that they cannot be reduced to anything basic within the confinement of space–time and causality. For example one cannot ascribe either spatiality or causality to human thoughts, dreams etc. Does one need a framework that transcends spatio–temporality to incorporate such non–physical ‘events’? Or can one explain them by using BE? The following views can be taken depending on one’s belief [Samal (1999)]:

• Modified BE, or $M(BE)$

What could be the basic substratum of these non–physical entities? Could they be understood in terms of any suitably modified physical substratum? At the classical level one might think of reducing them to ‘events’ which, unlike the physical events, do not have any reference to spatiality. Attempts have been made [Samal (1999)] to understand the non–physical entities like thoughts and dreams in terms of temporal events and correlation between them. Although such an approach may yield the kinematics of these non–physical entities, it is not clear how their dynamics i.e., evolution etc., can be understood in terms of temporal component alone without any external spatial input, when in the first place they have arose from perceptions that are meaningful only
in the context of spatio–temporality?. Secondly, it is not clear why the ‘mental events’ constructed after dropping the spatiality should require new set of laws that are different from the usual physical laws. At the quantum level one might try to have a suitable modification of the wave \( \psi \)-function to incorporate these non–physical entities. One may make the wave \( \psi \)-function depend on extra parameters [Samal (1999)], either physical or non–physical, to give it the extra degrees–of–freedom to mathematically include more information. But such a wave \( \psi \)-function bound to have severe problems at the level of interpretation. For example, if one includes an extra parameter called ‘meditation’ as a new degree of freedom apart from the usual ones, then how will one interpret squared modulus of the wave \( \psi \)-function? It will be certainly too crude to extend the Born rule to conclude that the squared modulus in this case will give the probability of finding a particle having certain meditation value. Hence this kind of modification will not be of much help except for the apparent satisfaction of being able to write an eigenvalue equation for dreams or emotions. This approach is certainly not capable of telling how the wave \( \psi \)-function is related to consciousness, let alone a mathematical equation for the evolution of consciousness.

If one accepts consciousness as a phenomenon that arises out of execution of processes then any suggested [Samal (2001)] new physical basis can be shown to be redundant. As we have concluded earlier, all such possible processes and their execution can be reduced to \( BE \) and spatio–temporal correlations among \( BE \) using a suitable definition of communication.

Hence to incorporate non–physical entities as some kind of information one has to modify the \( BE \) in a subtle way. Schematically \( M(BE) = BE \otimes X \), where \( \otimes \) stands for a yet unknown operation and \( X \) stands for fundamental substratum of non–physical information. \( X \) has to be different from \( BE \); otherwise it could be reducible to \( BE \) and then there will be no spatio–temporal distinction between physical and non–physical information. But, how to find out what is \( X \)? Is it evident that the laws for \( M(BE) \) will be different from that for \( BE \)?

- **Give up \( BE \)**

One could believe that it is the ‘Qualia’ that constitutes consciousness and hence consciousness has to be understood at a phenomenological level without dissecting it into \( BE \) or \( M(BE) \). One would note that consciousness mainly consists of three phenomenological processes that
can be roughly put as retentive, reflective and creative. But keeping the
tremendous progress of our physical sciences and their utility to neuro-
siences in view, it is not unreasonable to expect that all these three
phenomenological processes, involving both human as well as animal
can be understood one day in terms of $M(BE)$.

**Platonic** $BE$

It has been suggested [Samal (1999)] that consciousness could be like
mathematics in the sense that although it is needed to comprehend the
physical reality, in itself it is not ‘real’.

The ‘reality’ of mathematics is a controversial issue that brings in the
old debate between the realists and the constructivists whether a math-
ematical truth is ‘a discovery’ or ‘an invention’ of the human mind? Should one consider the physical laws based on mathematical truth as
real or not?. The realist’s stand of attributing a Platonic existence
to the mathematical truth is a matter of pure faith unless one tries
to get the guidance from the knowledge of the physical world. It is
doubtful whether our knowledge of physical sciences provides support
for the realist’s view if one considers the challenge to ‘realism’ in physi-
cal sciences by the quantum world–view, which has been substantiated
in recent past by experiments [Aspect *et al*. (1982)] that violate Bell’s
inequalities.

Even if one accepts the Platonic world of mathematical forms, this no
way makes consciousness non-existent or unreal. Rather the very fact
that truth of such a platonic world of mathematics yields to the human
understanding as much as that of a physical world makes consciousness
all the more profound in its existence.

### 7.1.4.2 Manipulation of Consciousness

Can consciousness be manipulated in a controlled manner? Experience tells
us how difficult it is to control the thoughts and how improbable it is to
control the dreams. We discuss below few methods prescribed by western
psychoanalysis and oriental philosophies regarding the manipulation of
consciousness [Samal (1999)]. Is there a lesson for modern science to learn
from these methods?

**Self**

The subject of ‘self’ is usually considered to belong to an ‘internal space’
in contrast to the external space where we deal with others. We will consider
the following two cases here:

- **Auto–suggestions**
  
  There have been evidences that by auto–suggestions one can control one’s feelings like pain and pleasure. Can one cure oneself of diseases of physical origin by auto–suggestions? This requires further investigations.

- **Yoga and other oriental methods**
  
  The eight–fold Yoga of Patanjali is perhaps the most ancient method prescribed [Iyengar (1981)] to control one’s thought and to direct it in a controlled manner. But it requires certain control over body and emotions before one aspires to gain control over mind. In particular it lays great stress on ‘breath control’ (pranayama) as a means to relax the body and to still the mind. In its later stages it provides systematic methods to acquire concentration and to prolong concentration on an object or a thought.

  After this attainment one can reach a stage where one’s awareness of self and the surrounding is at its best. Then in its last stage, Yoga prescribes one’s acute awareness to be decontextualized [Samal (1999)] from all perceptions limited by spatio–temporality and thus to reach a pinnacle called (samadhi) where one attains an understanding of everything and has no doubts. In this sense the Yogic philosophy believes that pure consciousness transcends all perceptions and awareness. It is difficult to understand this on the basis of day to day experience. Why does one need to sharpen one’s awareness to its extreme if one is finally going to abandon its use? How does abandoning one’s sharpened awareness help in attaining a realisation that transcends spatio–temporality? Can any one realise anything that is beyond the space, time and causality? What is the purpose of such a consciousness that lies beyond the confinement of space and time?

**Non–Self**

The Non–Self belongs to an external world consisting of others, both living and non–living. In the following we discuss whether one can direct one’s consciousness towards others such that one can affect their behavior [Samal (1999), Samal (2001)].
Quantum Leap

- Hypnosis, ESP, and Paranormal

It is a well-known fact that it is possible to hypnotize a person and then to make contact with his/her subconscious mind. Where does this subconscious lie? What is its relation to the conscious mind? The efficacy of the method of hypnosis in curing people of deep-rooted psychological problems tells us that we are yet to understand the dynamics of the human brain fully.

The field of Para-Psychology deals with ‘phenomena’ like Extra Sensory Perception (ESP) and telepathy etc. where one can direct one’s consciousness to gain insight into future or to influence others mind. It is not possible to explain these on the basis of the known laws of the world. It has been claimed that under hypnosis a subject could vividly recollect incidents from the previous lives including near-death and death experiences which is independent of spatio-temporality. Then, it is not clear, why most of these experiences are related to past? If these phenomena are truly independent of space and time, then studies should be made to find out if anybody under hypnosis can predict his/her own death, an event that can be easily verifiable in due course of time, unlike the recollections of past-life [Samal (1999)].

Can mind influence matter belonging to outside of the body? The studies dubbed as Psycho-Kinesis (PK) have been conducted to investigate the ‘suspect’ interaction of the human mind with various material objects such as cards, dice, simple pendulum etc. An excellent historical overview of such studies leading upto the modern era is available as a review paper, titled “The Persistent Paradox of Psychic Phenomena: An Engineering Perspective,” by Robert Jahn of Princeton University, published in Proc. IEEE (Feb. 1982).

The Princeton Engineering Anomalies Research (PEAR) programme of the Department of Applied Sciences and Engineering, Princeton University, has recently developed and patented a ‘Field REG’ (Field Random Event Generator) device which is basically a portable notebook computer with a built-in truly random number generator (based on a microelectronic device such as a shot noise resistor or a solid-state diode) and requisite software for on-line data processing and display, specifically tailored for conducting ‘mind-machine interaction’ studies. After performing large number of systematic experiments over the last two decades, the PEAR group has reported [Samal (1999)] the existence of such a consciousness related mind-machine interaction in the case of ‘truly random devices’. They attribute it to a ‘Consciousness Field
Effect'. They have also reported that deterministic random number sequences such as those generated by mathematical algorithm or pseudo–random generators do not show any consciousness related anomalous behavior. Another curious finding is that ‘intense emotional resonance’ generates the effect whereas ‘intense intellectual resonance’ does not. It is also not clear what is the strength of the ‘consciousness field’ in comparison to all the four known basic force fields of nature.

One should not reject outright any phenomenon that cannot be explained by the known basic laws of nature. Because each such phenomenon holds the key to extend the boundary of our knowledge further. But before accepting these effects one should filter them through the rigors of scientific methodology. In particular, the following questions can be asked [Samal (1999)]:

- Why are these events rare and not repeatable?
- How does one make sure that these effects are not manifestations of yet unknown facets of the known forces?
- Why is it necessary to have truly random processes? How does one make sure that these are not merely statistical artifacts?

If the above effects survive the scrutiny of the above questions (or similar ones) then they will open up the doors to a new world not yet known to science. In such a case how does one accommodate them within the existing framework of scientific methods? If these effects are confirmed beyond doubt, then one has to explore the possibility that at the fundamental level of nature, the laws are either different from the known physical laws or there is a need to complement the known physical laws with a set of non–physical laws. In such a situation, these ‘suspect’ phenomena might provide us with the valuable clue for modifying \( BE \) to get \( M(BE) \) that is the basis of everything including both physical and mental.

7.1.4.3 *A Need for Paradigm–Shift*

Although reductionist approach can provide us with valuable clues regarding the attributes of consciousness, it is the holistic approach that can only explain consciousness. But the dualism of Descartes that treats physical and mental processes in a mutually exclusive manner will not suffice for understanding consciousness unless it makes an appropriate use of complementarity for mental and physical events which is analogous to the com-
plementarity evident in the quantum world.

Where does the brain end and the mind begin? Brain is the physical means to acquire and to retain the information for the mind to process them to find a ‘meaning’ or a ‘structure’ which we call ‘understanding’ that is attributed to consciousness. Whereas attributes of consciousness can be reduced to $BE$ (or to $M(BE)$), the holistic process of consciousness can only be understood in terms of ‘quantum communication’, where ‘communication’ has an appropriate meaning. Maybe one has to look for such a suitable definition of communication at the level of ‘quantum vacuum’ [Samal (1999)].

7.1.4.4 Quantum Synapses

In 1986 Eccles [Eccles (1986)] offered plausible arguments for mental events causing neural events via the mechanism of wave function collapse. Conventional operations of the synapses depend on the operation of “ultimate synaptic units” called “boutons.” Eccles states that, “these synaptic boutons, when excited by an all-or-nothing nerve impulse, deliver the total content of a single synaptic vesicle, not regularly, but probabilistically.” Eccles points out that refined physiological analysis of the synapse shows that the effective structure of each bouton is a para-crystalline presynaptic vesicular grid with about 50 vesicles. The existence of such a crystalline structure is suggestive of quantum physical laws in operation in that the spacing and structure are suggestive of crystalline structure in common substances.

Eccles focused attention on these para-crystalline grids as the targets for non-material events. He showed how the probability field of quantum mechanics which carries neither mass nor energy, can nevertheless be envisioned as exerting effective action at these microsites. In the event of a sudden change in the probability field brought on by the observation of a complementary observable, there would be a change in the probability of emission of one or more of the vesicles.

Eccles concluded that calculations based on the Heisenberg uncertainty principle show that the probabilistic emission of a vesicle from the para-crystalline presynaptic grid could conceivably be modified by mental intention in the same manner that mental intention modifies a quantum wave function. Although my conclusions were based on the operations of protein gating molecules in the neural wall, I came to a similar conclusion: mental events stimulate neural events through sudden changes in the quantum
physical probability field and the timing of these events could be governed by quantum mechanical consideration.

7.1.4.5 Quantum Timing of Conscious Experience

Quantum systems in the time interval between two events, so-called two-time observables (TTO), are known to behave in a manner quite different from expectations based on initial value quantum mechanics. According to the transactional interpretation (TI) of quantum physics, wave functions can be pictured as offer and echo waves – the offer wave passing from an initial event, $i$, to a future event, $a$, and the echo wave, the complex conjugate of the offer wave, passing from $a$ back in time toward $i$. TTO and the TI have been used to explain certain quantum physical temporal anomalies, such as non-locality, contra-factuality, and future-to-present causation as explicitly shown in Wheeler’s delayed choice experiment, an experiment wherein the history of the objects under scrutiny are not determined until the final observation. Experimental evidence involving neurological functioning and subjective awareness indicates the presence of the same anomalies [Wolf (1998)].

7.1.5 Quantum Brain

7.1.5.1 Biochemistry of Microtubules

Recent developments/efforts to understand aspects of the brain function at the sub-neural level are discussed in [Nanopoulos (1995)]. Microtubules (MTs), protein polymers constructing the cytoskeleton of a neuron, participate in a wide variety of dynamical processes in the cell. Of special interest for this subsection is the MTs participation in bioinformation processes such as learning and memory, by possessing a well-known binary error-correcting code $[K_1(13,2^6,5)]$ with 64 words. In fact, MTs and DNA/RNA are unique cell structures that possess a code system. It seems that the MTs’ code system is strongly related to a kind of mental code in the following sense. The MTs’ periodic para-crystalline structure make them able to support a superposition of coherent quantum states, as it has been recently conjectured by Hameroff and Penrose [Hameroff and Penrose (1996)], representing an external or mental order, for sufficient time needed for efficient quantum computing [Ivancevic and Ivancevic (2007c); Ivancevic and Ivancevic (2007d)].

Living organisms are collective assemblies of cells which contain col-
lective assemblies of organized material, including membranes, organelles, nuclei, and the cytoplasm, the bulk interior medium of living cells. Dynamic rearrangements of the cytoplasm within eucaryotic cells, the cells of all animals and almost all plants on Earth, account for their changing shape, movement, etc. This extremely important cytoplasmic structural and dynamical organization is due to the presence of networks of interconnected protein polymers, which are referred to as the cytoskeleton due to their bone–like structure \cite{hameroff1996} \cite{dustin1984}. The cytoskeleton consists of MT’s, actin microfilaments, intermediate filaments and an organizing complex, the centrosome with its chief component the centriole, built from two bundles of microtubules in a separated T shape. Parallel–arrayed MTs are interconnected by cross–bridging proteins (MT–Associated Proteins: MAPs) to other MTs, organelle filaments and membranes to form dynamic networks \cite{hameroff1996} \cite{dustin1984}. MAPs may be contractile, structural, or enzymatic. A very important role is played by contractile MAPs, like dynein and kinesin, through their participation in cell movements as well as in intraneural, or axoplasmic transport which moves material and thus is of fundamental importance for the maintenance and regulation of synapses (see, e.g., \cite{eccles1986}). The structural bridges formed by MAPs stabilize MTs and prevent their disassembly. The MT–MAP ‘complexes’ or cytoskeletal networks determine the cell architecture and dynamic functions, such as mitosis, or cell division, growth, differentiation, movement, and for us here the very crucial, synapse formation and function, all essential to the living state. It is usually said that microtubules are ubiquitous through the entire biology \cite{hameroff1996} \cite{dustin1984}.

MTs are hollow cylinders comprised of an exterior surface of cross–section diameter 25 nm (1 nm = 10^{-9} meters) with 13 arrays (protofilaments) of protein dimers called tubulines \cite{dustin1984}. The interior of the cylinder, of cross–section diameter 14 nm, contains ordered water molecules, which implies the existence of an electric dipole moment and an electric field. The arrangement of the dimers is such that, if one ignores their size, they resemble triangular lattices on the MT surface. Each dimer consists of two hydrophobic protein pockets, and has an unpaired electron. There are two possible positions of the electron, called \( \alpha \) and \( \beta \) conformations. When the electron is in the \( \beta \)–conformation there is a 29° distortion of the electric dipole moment as compared to the \( \alpha \) conformation.

In standard models for the simulation of the MT dynamics \cite{satari},
et al. (1993) Satarić et al. (1998), the ‘physical’ DOF – relevant for the description of the energy transfer – is the projection of the electric dipole moment on the longitudinal symmetry axis (x-axis) of the MT cylinder. The 29° distortion of the β-conformation leads to a displacement $u_n$ along the $x$-axis, which is thus the relevant physical DOF.

There has been speculation for quite some time that MTs are involved in information processing: it has been shown that the particular geometrical arrangement (packing) of the tubulin protofilaments obeys an error-correcting mathematical code known as the $K_2(13, 2^9, 5)$-code [Koruga et al. (1993)]. Error correcting codes are also used in classical computers to protect against errors while in quantum computers special error correcting algorithms are used to protect against errors by preserving quantum coherence among qubits.

Information processing occurs via interactions among the MT protofilament chains. The system may be considered as similar to a model of interacting Ising chains on a triangular lattice, the latter being defined on the plane stemming from filleting open and flattening the cylindrical surface of MT. Classically, the various dimers can occur in either α or β conformations. Each dimer is influenced by the neighboring dimers resulting in the possibility of a transition. This is the basis for classical information processing, which constitutes the picture of a (classical) cellular automaton.

7.1.5.2 Kink Soliton Model of MT–Dynamics

The quantum nature of an MT network results from the assumption that each dimer finds itself in a superposition of α and β conformations. Viewed as a two-state quantum mechanical system, the MT tubulin dimers couple to conformational changes with $10^{-9} - 10^{-11}$ sec transitions, corresponding to an angular frequency $\omega \sim O(10^{10}) - O(10^{12})$ Hz [Nanopoulos (1995)].

The quantum computer character of the MT network [Penrose (1967)] results from the assumption that each dimer finds itself in a superposition of α and β conformations [Hameroff (1987)]. There is a macroscopic coherent state among the various chains, which lasts for $O(1\text{ sec})$ and constitutes the ‘preconscious’ state [Nanopoulos (1995)]. The interaction of the chains with (non–critical stringy) quantum gravity, then, induces self–collapse of the wave function of the coherent MT network, resulting in quantum computation.

In [Ellis et al. (1992); Ellis et al. (1999); Mavromatos and Nanopoulos (1995a); Mavromatos and Nanopoulos (1995b); Nanopoulos (1995)] the
authors assumed that the collapse occurs mainly due to the interaction of each chain with quantum gravity, the interaction from neighboring chains being taken into account by including mean–field interaction terms in the dynamics of the displacement field of each chain. This amounts to a modification of the effective potential by anharmonic oscillator terms. Thus, the effective system under study is 2D, possessing one space and one time coordinate.

Let \( u_n \) be the displacement field of the \( n \)th dimer in a MT chain. The continuous approximation proves sufficient for the study of phenomena associated with energy transfer in biological cells, and this implies that one can make the replacement

\[
\begin{align*}
  u_n &\rightarrow u(x,t), \quad (7.4)
\end{align*}
\]

with \( x \) a spatial coordinate along the longitudinal symmetry axis of the MT. There is a time variable \( t \) due to fluctuations of the displacements \( u(x) \) as a result of the dipole oscillations in the dimers.

The effects of the neighboring dimers (including neighboring chains) can be phenomenologically accounted for by an effective potential \( V(u) \). In the kink–soliton model\(^{30}\) of ref. Satarić et al. (1993); Satarić et al. (1998) a double–well potential was used, leading to a classical kink solution for the \( u(x,t) \) field. More complicated interactions are allowed in the picture of Ellis et al., where more generic polynomial potentials have been considered.

The effects of the surrounding water molecules can be summarized by a viscous force term that damps out the dimer oscillations,

\[
F = -\gamma \partial_t u, \quad (7.5)
\]

with \( \gamma \) determined phenomenologically at this stage. This friction should be viewed as an environmental effect, which however does not lead to energy dissipation, as a result of the non–trivial solitonic structure of the ground–state and the non–zero constant force due to the electric field. This is a well known result, directly relevant to energy transfer in biological systems.

In mathematical terms the effective equation of motion for the relevant field DOF \( u(x,t) \) reads:

\[
\begin{align*}
  u''(\xi) + \rho u'(\xi) &\rightarrow P(u), \quad (7.6)
\end{align*}
\]

where \( \xi = x - vt, \quad u'(\xi) = du/d\xi, \quad v \) is the velocity of the soliton, \( \rho \propto \gamma \)

\(^{30}\)Recall that kinks are solitary (non–dispersive) waves arising in various 1D (bio)physical systems.
\[ P(u) \text{ is a polynomial in } u, \text{ of a certain degree, stemming from the variations of the potential } V(u) \text{ describing interactions among the MT chains. In the mathematical literature there has been a classification of solutions of equations of this form. For certain forms of the potential the solutions include kink solitons that may be responsible for dissipation-free energy transfer in biological cells:} \]

\[ u(x, t) \sim c_1 \left( \tanh[c_2(x - vt)] + c_3 \right), \quad (7.7) \]

where \( c_1, c_2, c_3 \) are constants depending on the parameters of the dimer lattice model. For the form of the potential assumed in the model of Satarić et al. (1993); Satarić et al. (1998) there are solitons of the form \( u(x, t) = c'_1 + \frac{c'_2 - c'_1}{1 + e^{c'_1(x - vt)}} \), where again \( c'_i, i = 1, \ldots, 3 \) are appropriate constants.

A semiclassical quantization of such solitonic states has been considered by Ellis et al. The result of such a quantization yields a modified soliton equation for the (quantum corrected) field \( u_q(x, t) \) [Tsue and Fujiwara (1991)]

\[ \partial_t^2 u_q(x, t) - \partial_x^2 u_q(x, t) + M^{(1)}[u_q(x, t)] = 0, \quad (7.8) \]

with the notation

\[ M^{(n)} = e^{\frac{1}{2}(G(x, y, t) - G_0(x, y))} \left. \frac{d^n}{dz^n} U^{(n)}(z) \right|_{z = u_q(x, t)}, \quad U^{(n)} \equiv \frac{d^n U}{dz^n}. \]

The quantity \( U \) denotes the potential of the original soliton Hamiltonian, and \( G(x, y, t) \) is a bilocal field that describes quantum corrections due to the modified boson field around the soliton. The quantities \( M^{(n)} \) carry information about the quantum corrections. For the kink soliton (7.7) the quantum corrections (7.8) have been calculated explicitly in Tsue and Fujiwara (1991), thereby providing us with a concrete example of a large-scale quantum coherent state.

A typical propagation velocity of the kink solitons (e.g., in the model of Satarić et al. (1993); Satarić et al. (1998)) is \( v \sim 2 \text{ m/sec} \), although, models with \( v \sim 20 \text{ m/sec} \) have also been considered. This implies that, for moderately long microtubules of length \( L \sim 10^{-6} \text{ m} \), such kinks transport energy without dissipation in

\[ t_F \sim 5 \times 10^{-7} \text{ s.} \quad (7.9) \]

Such time scales are comparable to, or smaller in magnitude than, the decoherence time scale of the above-described coherent (solitonic) states.
quantum mechanical phenomena may then be responsible for frictionless energy (and signal) transfer across microtubular arrangements in the cell [Nanopoulos (1995)].

7.1.5.3 Liouville–Hamiltonian Neurodynamics

Recall that neurodynamics has its physical behavior both on the macroscopic, classical, inter–neuronal level, and on the microscopic, quantum, intra–neuronal level. On the macroscopic level, various models of neural networks (NNs, for short) have been proposed as goal–oriented models of the specific neural functions, like for instance, function–approximation, pattern–recognition, classification, or control (see, e.g., [Haykin (1994)]). In the physically–based, Hopfield–type models of NNs [Hopfield (1982); Hopfield (1984)] the information is stored as a content–addressable memory in which synaptic strengths are modified after the Hebbian rule (see [Hebb (1949)]). Its retrieval is made when the network with the symmetric couplings works as the point–attractor with the fixed points. Analysis of both activation and learning dynamics of Hopfield–Hebbian NNs using the techniques of statistical mechanics [Domany et al. (1991)] gives us with the most important information of storage capacity, role of noise and recall performance.

On the other hand, on the general microscopic intra–cellular level, energy transfer across the cells, without dissipation, had been first conjectured to occur in biological matter by [Fröhlich and Kremer (1983)]. The phenomenon conjectured by them was based on their 1D superconductivity model: in 1D electron systems with holes, the formation of solitonic structures due to electron–hole pairing results in the transfer of electric current without dissipation. In a similar manner, Fröhlich and Kremer conjectured that energy in biological matter could be transferred without dissipation, if appropriate solitonic structures are formed inside the cells. This idea has lead theorists to construct various models for the energy transfer across the cell, based on the formation of kink classical solutions (see [Satarić et al. (1993); Satarić et al. (1998)].

The interior of living cells is structurally and dynamically organized by cytoskeletons, i.e., networks of protein polymers. Of these structures, microtubules (MTs, for short) appear to be the most fundamental (see [Dustin (1984)]). Their dynamics has been studied by a number of authors in connection with the mechanism responsible for dissipation–free energy transfer. Hameroff and Penrose [Hameroff (1987)] have con-
jectured another fundamental role for the MTs, namely being responsible for quantum computations in the human neurons. [Penrose (1967); Penrose (1994); Penrose (1997)] further argued that the latter is associated with certain aspects of quantum theory that are believed to occur in the cytoskeleton MTs, in particular quantum superposition and subsequent collapse of the wave function of coherent MT networks. These ideas have been elaborated by [Mavromatos and Nanopoulos (1995a); Mavromatos and Nanopoulos (1995b); Nanopoulos (1995)], based on the quantum-gravity EMN-language of [Ellis et al. (1992); Ellis et al. (1999)] where MTs have been physically modelled as non-critical (SUSY) bosonic strings. It has been suggested that the neural MTs are the microsites for the emergence of stable, macroscopic quantum coherent states, identifiable with the preconscious states; stringy-quantum space-time effects trigger an organized collapse of the coherent states down to a specific or conscious state. More recently, [Tabony et al. (1999)] have presented the evidence for biological self-organization and pattern formation during embryogenesis.

Now, we have two space-time biophysical scales of neurodynamics. Naturally the question arises: are these two scales somehow inter-related, is there a space-time self-similarity between them?

The purpose of this subsection is to prove the formal positive answer to the self-similarity question. We try to describe neurodynamics on both physical levels by the unique form of a single equation, namely open Liouville equation: NN-dynamics using its classical form, and MT-dynamics using its quantum form in the Heisenberg picture. If this formulation is consistent, that would prove the existence of the formal neurobiological space-time self-similarity.

**Hamiltonian Framework**

Suppose that on the macroscopic NN-level we have a conservative Hamiltonian system acting in a $2N$-dimensional symplectic phase space $T^*Q = \{q^i(t), p_i(t)\}, i = 1 \ldots N$ (which is the cotangent bundle of the NN-configuration manifold $Q = \{q^i\}$), with a Hamiltonian function $H = H(q^i, p_i, t) : T^*Q \times \mathbb{R} \to \mathbb{R}$ (see [Ivancevic and Ivancevic (2005); Ivancevic and Ivancevic (2006b)]). The conservative dynamics is defined by classical Hamilton’s canonical equations:

\begin{align}
 \dot{q}^i &= \partial_{p_i} H \quad \text{contravariant velocity equation}, \\
 \dot{p}_i &= -\partial_{q^i} H \quad \text{covariant force equation},
\end{align}

(7.10)
(here and henceforth overdot denotes the total time derivative). Within the framework of the conservative Hamiltonian system (7.10) we can apply the formalism of classical Poisson brackets: for any two functions \( A = A(q^i, p_i, t) \) and \( B = B(q^i, p_i, t) \) their Poisson bracket is (using the summation convention) defined as \cite{Ivancevic and Ivancevic (2006b)}

\[
[A, B] = (\partial_{q^i} A \partial_{p_i} B - \partial_{p_i} A \partial_{q^i} B).
\]

**Conservative Classical System**

Any function \( A(q^i, p_i, t) \) is called a constant (or integral) of motion of the conservative system (7.10) if \cite{Ivancevic and Ivancevic (2007b)}

\[
\dot{A} \equiv \partial_t A + [A, H] = 0,
\]

(7.11)

For example, if \( A = \rho(q^i, p_i, t) \) is a density function of ensemble phase-points (or, a probability density to see a state \( x(t) = (q^i(t), p_i(t)) \) of ensemble at a moment \( t \)), then equation

\[
\partial_t \rho = -[\rho, H]
\]

(7.12)

represents the Liouville theorem, which is usually derived from the continuity equation

\[
\partial_t \rho + \text{div}(\rho \dot{x}) = 0.
\]

**Conservative Quantum System**

We perform the formal quantization of the conservative equation (7.12) in the Heisenberg picture: all variables become Hermitian operators (denoted by ‘\( \hat{\,} \)’), the symplectic phase space \( T^*Q = \{q^i, p_i\} \) becomes the Hilbert state space \( \mathcal{H} = \mathcal{H}_{q^1} \otimes \cdots \otimes \mathcal{H}_{q^N} \) and \( \mathcal{H}_{p_i} = \mathcal{H}_{p_1} \otimes \cdots \otimes \mathcal{H}_{p_N} \), the classical Poisson bracket \( [,] \) becomes the quantum commutator \( [\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\,\
where $\hat{H} = \hat{H}(\hat{q}_i, \hat{p}_i, t)$ is the Hamiltonian evolution operator, while

$$\hat{\rho} = \sum_a P(a) |\Psi_a\rangle \langle \Psi_a|,$$

with $\text{Tr}(\hat{\rho}) = 1$

denotes the von Neumann density matrix operator, where each quantum state $|\Psi_a\rangle$ occurs with probability $P(a)$; $\hat{\rho} = \hat{\rho}(\hat{q}_i^*, \hat{p}_i, t)$ is closely related to another von Neumann concept: entropy

$$S = -\text{Tr}(\hat{\rho} [\ln \hat{\rho}]).$$

**Open Classical System**

We now move to the open (nonconservative) system: on the macroscopic NN-level the opening operation equals to the adding of a covariant vector of external (dissipative and/or motor) forces $F_i = F_i(q_i, p_i, t)$ to (the right-hand-side of) the covariant Hamilton’s force equation, so that Hamilton’s equations obtain the open (dissipative and/or forced) form [Ivancevic and Ivancevic (2005); Ivancevic and Ivancevic (2006b)]:

$$\dot{q}_i = \partial_{p_i} H, \quad \dot{p}_i = -\partial_{q_i} H + F_i.$$  \hspace{1cm} (7.15)

In the framework of the open Hamiltonian system (7.15) dynamics of any function $A(q_i, p_i, t)$ is defined by the open (dissipative and/or forced) evolution equation:

$$\partial_t A = -[A, H] + F_i [A, q_i], \quad ([A, q_i] = -\partial_{p_i} A).$$  \hspace{1cm} (7.16)

In particular, if $A = \rho(q_i, p_i, t)$ represents the density function of ensemble phase-points then its dynamics is given by the open (dissipative and/or forced) Liouville equation [Ivancevic and Ivancevic (2006b); Ivancevic and Ivancevic (2007b)]:

$$\partial_t \rho = -[\rho, H] + F_i [\rho, q_i].$$  \hspace{1cm} (7.17)

Equation (7.17) represents the open classical model of our microscopic NN-dynamics.
Continuous Neural Network Dynamics

The generalized NN-dynamics, including two special cases of graded response neurons (GRN) and coupled neural oscillators (CNO), can be presented in the form of a Langevin stochastic equation [Ivančević and Ivančević (2006b)]

\[ \dot{\sigma}_i = f_i + \eta_i(t), \]  

(7.18)

where \( \sigma_i = \sigma_i(t) \) are the continual neuronal variables of \( i \)-th neurons (representing either membrane action potentials in case of GRN, or oscillator phases in case of CNO); \( J_{ij} \) are individual synaptic weights; \( f_i = f_i(\sigma_i, J_{ij}) \) are the deterministic forces (given, in GRN-case, by \( f_i = \sum_j J_{ij} \tanh[\gamma \sigma_j] - \sigma_i + \theta_i, \) with \( \gamma > 0 \)

and with the \( \theta_i \) representing injected currents, and in CNO-case, by

\[ f_i = \sum_j J_{ij} \sin(\sigma_j - \sigma_i) + \omega_i, \]

with \( \omega_i \) representing the natural frequencies of the individual oscillators); the noise variables are given as

\[ \eta_i(t) = \lim_{\Delta \to 0} \zeta_i(t) \sqrt{2T/\Delta}, \]

where \( \zeta_i(t) \) denote uncorrelated Gaussian distributed random forces and the parameter \( T \) controls the amount of noise in the system, ranging from \( T = 0 \) (deterministic dynamics) to \( T = \infty \) (completely random dynamics).

More convenient description of the neural random process (7.18) is provided by the Fokker-Planck equation describing the time evolution of the probability density \( P(\sigma_i) \) [Ivančević and Ivančević (2006b)]

\[ \partial_t P(\sigma_i) = - \sum_i \partial_{\sigma_i} [f_i P(\sigma_i)] + T \sum_i \partial_{\sigma_i^2} P(\sigma_i). \]  

(7.19)

Now, in the case of deterministic dynamics \( T = 0 \), equation (7.19) can be easily put into the form of the conservative Liouville equation (7.12), by making the substitutions:

\[ P(\sigma_i) \rightarrow \rho, f_i = \dot{\sigma}_i, \quad \text{and} \quad [\rho, H] = \text{div}(\rho \dot{\sigma}_i) \equiv \sum_i \partial_{\sigma_i} (\rho \dot{\sigma}_i), \]
where $H = H(\sigma_i, J_{ij})$. Further, we can formally identify the stochastic forces, i.e., the second-order noise-term $T \sum_i \partial_{\sigma_i^2} \rho$ with $F^i[\rho, \sigma_i]$, to get the open Liouville equation (7.17).

Therefore, on the NN-level deterministic dynamics corresponds to the conservative system (7.12). Inclusion of stochastic forces corresponds to the system opening (7.17), implying the macroscopic arrow of time.

Open Quantum System

By formal quantization of equation (7.17), in the same way as in paragraph II., we obtain the quantum open Liouville equation [Ivancevic and Ivancevic (2006b)]

$$\partial_t \hat{\rho} = i\{\hat{\rho}, \hat{H}\} - i\hat{F}_i\{\hat{\rho}, \hat{q}^i\},$$

(7.20)

where $\hat{F}_i = \hat{F}_i(\hat{q}^i, \hat{p}_i, t)$ represents the covariant quantum operator of external friction forces in the Hilbert state space $\mathcal{H} = \mathcal{H}_{\hat{q}_i} \otimes \mathcal{H}_{\hat{p}_i}$.

Equation (7.20) represents the open quantum-friction model of our microscopic MT-dynamics. Its system-independent properties are [Ellis et al. (1992); Ellis et al. (1999); Mavromatos and Nanopoulos (1995a); Mavromatos and Nanopoulos (1995b); Nanopoulos (1995)]:

1. Conservation of probability $P$

$$\partial_t P = \partial_t [\text{Tr}(\hat{\rho})] = 0.$$

2. Conservation of energy $E$, on the average

$$\partial_t \langle\langle E\rangle\rangle = \partial_t [\text{Tr}(\hat{\rho} E)] = 0.$$

3. Monotonic increase in entropy

$$\partial_t S = \partial_t [-\text{Tr}(\hat{\rho} \ln \hat{\rho})] \geq 0,$$

and thus automatically and naturally implies a microscopic arrow of time, so essential in realistic biophysics of neural processes.

Non–Critical Stringy MT–Dynamics

In EMN-language of non–critical (SUSY) bosonic strings, our MT-dynamics equation (7.20) reads

$$\partial_t \hat{\rho} = i\{\hat{\rho}, \hat{H}\} - i\hat{g}_{ij}\{\hat{\rho}, \hat{q}^i\}\hat{q}^j,$$

(7.21)
where the target–space density matrix $\hat{\rho}(\hat{q}_i, \hat{p}_i)$ is viewed as a function of coordinates $\hat{q}_i$ that parameterize the couplings of the generalized $\sigma$–models on the bosonic string world–sheet, and their conjugate momenta $\hat{p}_i$, while $\hat{g}_{ij} = \hat{g}_{ij}(\hat{q}_i)$ is the quantum operator of the \textit{positive definite metric} in the space of couplings. Therefore, the covariant quantum operator of external friction forces is in EMN–formulation given as $\hat{F}_i(\hat{q}_i, \hat{\dot{q}}_i) = \hat{g}_{ij}\hat{\dot{q}}_j$.

Equation (7.21) establishes the conditions under which a large–scale coherent state appearing in the MT–network, which can be considered responsible for loss–free energy transfer along the tubulins.

\section*{7.1.5.4 Dissipative Quantum Brain Model}

The \textit{conservative brain} model was originally formulated within the framework of the quantum field theory (QFT) by [Ricciardi and Umezawa (1967)] and subsequently developed in [Stuart \textit{et al.} (1978); Stuart \textit{et al.} (1979); Jibu and Yasue (1995); Jibu \textit{et al.} (1996)]. The conservative brain model has been recently extended to the \textit{dissipative quantum dynamics} in the work of G. Vitiello and collaborators [Vitiello (2005); Alfinito and Vitiello (2000); Pessa and Vitiello (1999); Vitiello (2001); Pessa and Vitiello (2003); Pessa and Vitiello (2004)].

The motivations at the basis of the formulation of the quantum brain model by Umezawa and Ricciardi trace back to the laboratory observations leading Lashley to remark (in 1940) that “masses of excitations... within general fields of activity, without regard to particular nerve cells are involved in the determination of behavior” [Lashley (1942); Pribram (1991)]. In 1960’s, K. Pribram, also motivated by experimental observations, started to formulate his \textit{holographic hypothesis}. According to
W. Freeman [Freeman (1990); Freeman (1996); Freeman (2000)], “information appears indeed in such observations to be spatially uniform in much the way that the information density is uniform in a hologram”. While the activity of the single neuron is experimentally observed in form of discrete and stochastic pulse trains and point processes, the ‘macroscopic’ activity of large assembly of neurons appears to be spatially coherent and highly structured in phase and amplitude.

Motivated by such an experimental situation, Umezawa and Ricciardi formulated in [Ricciardi and Umezawa (1967)] the quantum brain model as a many-body physics problem, using the formalism of QFT with spontaneous breakdown of symmetry (which had been successfully tested in condensed matter experiments). Such a formalism provides the only available theoretical tool capable to describe long-range correlations such as the ones observed in the brain – presenting almost simultaneous responses in several regions to some external stimuli. The understanding of these long-range correlations in terms of modern biochemical and electrochemical processes is still lacking, which suggests that these responses could not be explained in terms of single neuron activity [Pribram (1971); Pribram (1991)].

Lagrangian dynamics in QFT is, in general, invariant under some group $G$ of continuous transformations, as proposed by the famous Noether theorem. Now, spontaneous symmetry breakdown, one of the cornerstones of Haken’s synergetics [Haken (1983); Haken (1993)], occurs when the minimum energy state (the ground, or vacuum, state) of the system is not invariant under the full group $G$, but under one of its subgroups. Then it can be shown [Itzykson and Zuber (1980); Umezawa (1993)] that collective modes, the so-called Nambu–Goldstone (NG) boson modes, are dynamically generated. Propagating over the whole system, these modes are the carriers of the long-range correlation, in which the order manifests itself as a global property dynamically generated. The long-range correlation modes are responsible for keeping the ordered pattern: they are coherently condensed in the ground state (similar to e.g., in the crystal case, where they keep the atoms trapped in their lattice sites). The long-range correlation thus forms a sort of net, extending over all the system volume, which traps the system components in the ordered pattern. This explains the “holistic” macroscopic collective behavior of the system components.

More precisely, according to the Goldstone theorem in QFT [Itzykson and Zuber (1980); Umezawa (1993)], the spontaneous breakdown of the
symmetry implies the existence of long–range correlation NG–modes in the ground state of the system. These modes are massless modes in the infinite volume limit, but they may acquire a finite, non-zero mass due to boundary or impurity effects [Alfinito et al. (2002)]. In the quantum brain model, these modes are called dipole–wave–quanta (DWQ). The density of their condensation in the ground states acts as a code classifying the state and the memory there recorded. States with different code values are unitarily inequivalent states, i.e., there is no unitary transformation relating states of different codes.

Now, in formulating a proper mathematical model of brain, the conservative dynamics is not realistic: we cannot avoid to take into consideration the dissipative character of brain dynamics, since the brain is an intrinsically open system, continuously interacting with the environment. As Vitiello observed in [Vitiello (2001); Pessa and Vitiello (2003); Pessa and Vitiello (2004)], the very same fact of “getting an information” introduces a partition in the time coordinate, so that one may distinguish between before “getting the information” (the past) and after “getting the information” (the future): the arrow of time is in this way introduced. ...“Now you know it!” is the familiar warning to mean that now, i.e. after having received a certain information, you are not the same person as before getting it. It has been shown that the psychological arrow of time (arising as an effect of memory recording) points in the same direction of the thermodynamical arrow of time (increasing entropy direction) and of the cosmological arrow of time (the expanding Universe direction) [Alfinito et al. (2000)].

The canonical quantization procedure of a dissipative system requires to include in the formalism also the system representing the environment (usually the heat bath) in which the system is embedded. One possible way to do that is to depict the environment as the time–reversal image of the system [Celeghini et al. (1992)]; the environment is thus described as the double of the system in the time–reversed dynamics (the system image in the mirror of time).

Within the framework of dissipative QFT, the brain system is described in terms of an infinite collection of damped harmonic oscillators $A_\kappa$ (the

---

31 We remark that the spontaneous breakdown of symmetry is possible since in QFT there exist infinitely many ground states or vacua which are physically distinct (technically speaking, they are “unitarily inequivalent”). In quantum mechanics (QM), on the contrary, all the vacua are physically equivalent and thus there cannot be symmetry breakdown.
simplest prototype of a dissipative system) representing the DWQ [Vitiello (2095)]. Now, the collection of damped harmonic oscillators is ruled by the Hamiltonian [Vitiello (2095); Celeghini et al. (1992)]
\[ H = H_0 + H_I, \]
with
\[ H_0 = \hbar \Omega \kappa (A_\kappa^\dagger A_\kappa - \tilde{A}_\kappa^\dagger \tilde{A}_\kappa), \]
\[ H_I = i\hbar \Gamma \kappa (A_\kappa^\dagger \tilde{A}_\kappa - A_\kappa \tilde{A}_\kappa^\dagger), \]
where \( \Omega \kappa \) is the frequency and \( \Gamma \kappa \) is the damping constant. The \( \tilde{A}_\kappa \) modes are the ‘time–reversed mirror image’ (i.e., the ‘mirror modes’) of the \( A_\kappa \) modes. They are the doubled modes, representing the environment modes, in such a way that \( \kappa \) generically labels their degrees–of–freedom. In particular, we consider the damped harmonic oscillator (DHO)
\[ m \ddot{x} + \gamma \dot{x} + \kappa x = 0, \tag{7.22} \]
as a simple prototype for dissipative systems (with intention that thus get results also apply to more general systems). The damped oscillator (7.22) is a non–Hamiltonian system and therefore the customary canonical quantization procedure cannot be followed. However, one can face the problem by resorting to well known tools such as the density matrix \( \rho \) and the Wigner function \( W = W(x, p, t) \).

Let us start with the special case of a conservative particle in the absence of friction \( \gamma \), with the standard Hamiltonian,
\[ H = -(\hbar \partial_x)^2/2m + V(x). \]
Recall (from the previous subsection) that the density matrix equation of motion, i.e., quantum Liouville equation, is given by
\[ i\hbar \dot{\rho} = [H, \rho]. \tag{7.23} \]
The density matrix function \( \rho \) is defined by
\[ \langle x + \frac{1}{2} y | \rho(t) | x - \frac{1}{2} y \rangle = \psi^*(x + \frac{1}{2} y, t) \psi(x - \frac{1}{2} y, t) \equiv W(x, y, t), \]
with the associated standard expression for the Wigner function (see, e.g., Ivancevic and Ivancevic (2007b)),
\[ W(p, x, t) = \frac{1}{2\pi \hbar} \int W(x, y, t) e^{(-i\frac{p}{\hbar} y)} dy. \]
Now, in the coordinate \( x \)–representation, by introducing the notation
\[ x_\pm = x \pm \frac{1}{2} y, \tag{7.24} \]
the Liouville equation (7.23) can be expanded as
\[ i\hbar \partial_t \langle x_+ | \rho(t) | x_- \rangle = \left\{ -\frac{\hbar^2}{2m} \left[ \partial_{x_+}^2 - \partial_{x_-}^2 \right] + [V(x_+) - V(x_-)] \right\} \langle x_+ | \rho(t) | x_- \rangle, \tag{7.25} \]
while the Wigner function \( W(p,x,t) \) is now given by
\[ i\hbar \partial_t W(x,y,t) = H_0 W(x,y,t), \tag{7.26} \]
with
\[ H_0 = \frac{1}{m} p_x p_y + V(x + \frac{1}{2} y) - V(x - \frac{1}{2} y), \]
and \( p_x = -i\hbar \partial_x, \quad p_y = -i\hbar \partial_y. \)
The new Hamiltonian \( H_0 \) (7.26) may be get from the corresponding Lagrangian
\[ L_0 = \frac{\dot{x} \dot{y}}{2} - V(x + \frac{1}{2} y) + V(x - \frac{1}{2} y). \tag{7.27} \]
In this way, Vitiello concluded that the density matrix and the Wigner function formalism required, even in the conservative case (with zero mechanical resistance \( \gamma \)), the introduction of a ‘doubled’ set of coordinates, \( x_{\pm} \), or, alternatively, \( x \) and \( y \). One may understand this as related to the introduction of the ‘couple’ of indices necessary to label the density matrix elements (7.25).

Let us now consider the case of the particle interacting with a thermal bath at temperature \( T \). Let \( f \) denote the random force on the particle at the position \( x \) due to the bath. The interaction Hamiltonian between the bath and the particle is written as
\[ H_{\text{int}} = -fx. \tag{7.28} \]
Now, in the Feynman–Vernon formalism (see [Feynman (1972)]), the effective action \( A[x,y] \) for the particle is given by
\[ A[x,y] = \int_{t_i}^{t_f} L_0(\dot{x}, \dot{y}, x, y) dt + I[x,y], \]
with \( L_0 \) defined by (7.27) and
\[ e^{\frac{i}{\hbar} I[x,y]} = \langle e^{-\frac{i}{\hbar} \int_{t_i}^{t_f} f(t)x_-(t) dt} e^{\frac{i}{\hbar} \int_{t_i}^{t_f} f(t)x_+(t) dt} \rangle_+, \tag{7.29} \]
where the symbol \( \langle . \rangle \) denotes the average with respect to the thermal bath; \( \langle . \rangle_+ \) and \( \langle . \rangle_- \) denote time ordering and anti–time ordering, respectively; the coordinates \( x_{\pm} \) are defined as in (7.24). If the interaction between the
bath and the coordinate \( x \) were turned off, then the operator \( f \) of the bath would develop in time according to

\[
f(t) = e^{iH_\gamma t/\hbar} \left[ e^{-iH_\gamma t/\hbar} f \right],
\]

where \( H_\gamma \) is the Hamiltonian of the isolated bath (decoupled from the coordinate \( x \)). \( f(t) \) is then the force operator of the bath to be used in (7.29).

The interaction \( I[x,y] \) between the bath and the particle has been evaluated in [Srivastava et al. (1995)] for a linear passive damping due to thermal bath by following Feynman–Vernon and Schwinger. The final result from [Srivastava et al. (1995)] is:

\[
I[x,y] = \frac{1}{2} \int_{t_1}^{t_f} dt \left[ x(t) F_{y}^{\text{ret}}(t) + y(t) F_{x}^{\text{adv}}(t) \right]
+ \frac{i}{2\hbar} \int_{t_1}^{t_f} \int_{t_1}^{t_f} dtds N(t-s) y(t) y(s),
\]

where the retarded force on \( y \), \( F_{y}^{\text{ret}} \), and the advanced force on \( x \), \( F_{x}^{\text{adv}} \), are given in terms of the retarded and advanced Green functions \( G_{\text{ret}}(t-s) \) and \( G_{\text{adv}}(t-s) \) by

\[
F_{y}^{\text{ret}}(t) = \int_{t_1}^{t_f} ds \, G_{\text{ret}}(t-s) y(s), \quad F_{x}^{\text{adv}}(t) = \int_{t_1}^{t_f} ds \, G_{\text{adv}}(t-s) x(s),
\]

respectively. In (7.30), \( N(t-s) \) is the quantum noise in the fluctuating random force given by

\[
N(t-s) = \frac{1}{2} \langle f(t) f(s) + f(s) f(t) \rangle.
\]

The real and the imaginary part of the action are given respectively by

\[
\text{Re} \left( A[x,y] \right) = \int_{t_1}^{t_f} L \, dt,
\]

\[
L = m \ddot{x} - \left[ V(x + \frac{1}{2} y) - V(x - \frac{1}{2} y) \right] + \frac{1}{2} \left[ x F_{y}^{\text{ret}} + y F_{x}^{\text{adv}} \right], \quad (7.31)
\]

and

\[
\text{Im} \left( A[x,y] \right) = \frac{1}{2\hbar} \int_{t_1}^{t_f} \int_{t_1}^{t_f} N(t-s) y(t) y(s) \, dtds. \quad (7.32)
\]

Equations (7.30–7.32), are exact results for linear passive damping due to the bath. They show that in the classical limit ‘\( \hbar \to 0 \)’ nonzero \( y \) yields an
'unlikely process' in view of the large imaginary part of the action implicit in (7.32). Nonzero $y$, indeed, may lead to a negative real exponent in the evolution operator, which in the limit $\hbar \to 0$ may produce a negligible contribution to the probability amplitude. On the contrary, at quantum level nonzero $y$ accounts for quantum noise effects in the fluctuating random force in the system–environment coupling arising from the imaginary part of the action (see Srivastava et al. (1995)).

When in (7.31) we use $F_{y\text{ret}} = \gamma \dot{y}$ and $F_{x\text{adv}} = -\gamma \dot{x}$ we get,

$$L(\dot{x}, \dot{y}, x, y) = m\ddot{x}\dot{y} - V \left( x + \frac{1}{2}y \right) + V \left( x - \frac{1}{2}y \right) + \frac{\gamma}{2} (\dot{x}\dot{y} - \dot{y}\dot{x}). \quad (7.33)$$

By using

$$V \left( x \pm \frac{1}{2}y \right) = \frac{1}{2} \kappa (x \pm \frac{1}{2}y)^2$$

in (7.33), the DHO equation (7.22) and its complementary equation for the $y$ coordinate

$$m\ddot{y} - \gamma \dot{y} + \kappa y = 0. \quad (7.34)$$

are derived. The $y$–oscillator is the time–reversed image of the $x$–oscillator (7.22). From the manifolds of solutions to equations (7.22) and (7.34), we could choose those for which the $y$ coordinate is constrained to be zero, they simplify to

$$m\ddot{x} + \gamma \dot{x} + \kappa x = 0, \quad y = 0.$$ 

Thus we get the classical damped oscillator equation from a Lagrangian theory at the expense of introducing an ‘extra’ coordinate $y$, later constrained to vanish. Note that the constraint $y(t) = 0$ is not in violation of the equations of motion since it is a true solution to (7.22) and (7.34).

Therefore, the general scheme of the dissipative quantum brain model can be summarized as follows. The starting point is that the brain is permanently coupled to the environment. Of course, the specific details of such a coupling may be very intricate and changeable so that they are difficult to be measured and known. One possible strategy is to average the effects of the coupling and represent them, at some degree of accuracy, by means of some ‘effective’ interaction. Another possibility is to take into account the environmental influence on the brain by a suitable choice of
the brain vacuum state. Such a choice is triggered by the external input (breakdown of the symmetry), and it actually is the end point of the internal (spontaneous) dynamical process of the brain (self-organization). The chosen vacuum thus carries the *signature* (memory) of the reciprocal brain–environment influence at a given time under given boundary conditions. A change in the brain–environment reciprocal influence then would correspond to a change in the choice of the brain vacuum: the brain state evolution or ‘story’ is thus the story of the trade of the brain with the surrounding world. The theory should then provide the equations describing the brain evolution ‘through the vacua’, each vacuum for each instant of time of its history.

The brain evolution is thus similar to a time–ordered sequence of photograms: each photogram represents the ‘picture’ of the brain at a given instant of time. Putting together these photograms in ‘temporal order’ one gets a movie, i.e. the story (the evolution) of open brain, which includes the brain–environment interaction effects.

The evolution of a memory specified by a given code value, say \( \mathcal{N} \), can be then represented as a trajectory of given initial condition running over time–dependent vacuum states, denoted by \( |0(t)\rangle_{\mathcal{N}} \), each one minimizing the free energy functional. These trajectories are known to be classical trajectories in the infinite volume limit: transition from one representation to another inequivalent one would be strictly forbidden in a quantum dynamics.

Since we have now two–modes (i.e., non–tilde and tilde modes), the memory state \( |0(t)\rangle_{\mathcal{N}} \) turns out to be a two–mode coherent state. This is known to be an entangled state, i.e., it cannot be factorized into two single–mode states, the non–tilde and the tilde one. The physical meaning of such an entanglement between non-tilde and tilde modes is in the fact that the brain dynamics is permanently a dissipative dynamics. The entanglement, which is an unavoidable mathematical result of dissipation, represents the impossibility of cutting the links between the brain and the external world.\(^{32}\)

---

\(^{32}\)We remark that the entanglement is permanent in the large volume limit. Due to boundary effects, however, a unitary transformation could disentangle the tilde and non–tilde sectors: this may result in a pathological state for the brain. It is known that forced isolation of a subject produces pathological states of various kinds. We also observe that the tilde mode is not just a mathematical fiction. It corresponds to a real excitation mode (quasi–particle) of the brain arising as an effect of its interaction with the environment: the couples of non–tilde/tilde dwq quanta represent the correlation modes dynamically created in the brain as a response to the brain–environment reciprocal influence. It is the interaction between tilde and non–tilde modes that controls the irreversible time evolution of the brain: these collective modes are confined to live in the brain. They
In the dissipative brain model, noise and chaos turn out to be natural ingredients of the model. In particular, in the infinite volume limit the chaotic behavior of the trajectories in memory space may account for the high perceptive resolution in the recognition of the perceptual inputs. Indeed, small differences in the codes associated to external inputs may lead to diverging differences in the corresponding memory paths. On the other side, it also happens that codes differing only in a finite number of their components (in the momentum space) may easily be recognized as being the ‘same’ code, which makes possible that ‘almost similar’ inputs are recognized by the brain as ‘equal’ inputs (as in pattern recognition).

Therefore, the brain may be viewed as a complex system with (infinitely) many macroscopic configurations (the memory states). Dissipation is recognized to be the root of such a complexity.

7.1.5.5 QED Brain

In this subsection, mainly following [Stapp (1995)], we formulate a quantum electrodynamics brain model. Recall that quantum electrodynamics (extended to cover the magnetic properties of nuclei) is the theory that controls, as far as we know, the properties of the tissues and the aqueous (ionic) solutions that constitute our brains. This theory is our paradigm basic physical theory, and the one best understood by physicists. It describes accurately, as far as we know, the huge range of actual physical phenomena involving the materials encountered in daily life. It is also related to classical electrodynamics in a particularly beautiful and useful way.

In the low–energy regime of interest here it should be sufficient to consider just the classical part of the photon interaction defined in [Stapp (1983)]. Then the explicit expression for the unitary operator that describes the evolution from time \( t_1 \) to time \( t_2 \) of the quantum electromagnetic field in the presence of a set \( L = \{L_i\} \) of specified classical charged–particle trajectories, with trajectory \( L_i \) specified by the function \( x_i(t) \) and carrying charge \( e_i \), is [Stapp (1995)]

\[
U[L; t_2, t_1] = \exp(a^* \cdot J(L)) \exp(-J^*(L) \cdot a) \exp[-(J^*(L) \cdot J(L))/2],
\]

where, for any \( X \) and \( Y \),

\[
\langle X \cdot Y \rangle \equiv \int d^4k(2\pi)^{-4}2\pi\delta^+(k^2)X(k) \cdot Y(k),
\]

vanish as soon as the links between the brain and the environment are cut.
Quantum Mind

\[(X \cdot Y) \equiv \int d^4k (2\pi)^{-4} i(k^2 + i\epsilon)^{-1} X(k) \cdot Y(k),\]

and \(X \cdot Y = X^\mu Y_\mu = X^\mu Y_\mu\). Also,

\[J_\mu(L; k) \equiv \sum_i -ie_i \int_{L_i} dx_\mu \exp(ikx).\]

The integral along the trajectory \(L_i\) is

\[\int_{L_i} dx_\mu \exp(ikx) \equiv \int_{t_1}^{t_2} dt (dx_\mu(t)/dt) \exp(ikx).\]

The \(a^*(k)\) and \(a(k)\) are the photon creation and annihilation operators:

\[\{a(k), a^*(k')\} = (2\pi)^3 \delta^3(k - k')2k_0.\]

The operator \(U[L; t_2, t_1]\) acting on the photon vacuum state creates the coherent photon state that is the quantum–theoretic analog of the classical electromagnetic field generated by classical point particles moving on the set of trajectories \(L = \{L_i\}\) between times \(t_1\) and \(t_2\).

The \(U[L; t_2, t_1]\) can be decomposed into commuting contributions from the various values of \(k\). The general coherent state can be written [Stapp (1995)]

\[|q, p\rangle \equiv \exp i((q \cdot P) - (p \cdot Q))[0],\]

where \(|0\rangle\) is the photon vacuum state and

\[Q(k) = (a_k + a_k^\dagger)/\sqrt{2} \quad \text{and} \quad P(k) = i(a_k - a_k^\dagger)/\sqrt{2},\]

and \(q(k)\) and \(p(k)\) are two functions defined (and square integrable) on the mass shell \(k^2 = 0, k_0 \geq 0\). The inner product of two coherent states is

\[\langle q, p|q', p'\rangle = \exp -((q - q' \cdot q - q') + (p - p' \cdot p - p') + 2i(p - p' \cdot q + q'))/4.\]

There is a decomposition of unity

\[I = \prod d^4k (2\pi)^{-4} 2\pi \delta^+(k^2) \int dq_k dp_k/\pi\]

\[\times \exp(iq_k P_k - ip_k Q_k)|0_k\rangle \langle 0_k| \exp -(iq_k P_k - ip_k Q_k).\]
Quantum Leap

Here meaning can be given by quantizing in a box, so that the variable \( k \) is discretized. Equivalently,

\[
I = \int \mu(q, p) |q, p\rangle \langle q, p|
\]

where \( \mu(q, p) \) is the appropriate measure on the functions \( q(k) \) and \( p(k) \). Then if the state \( |\Psi\rangle \langle \Psi| \) were to jump to \( |q, p\rangle \langle q, p| \) with probability density \( \langle q, p|\Psi\rangle \langle \Psi|q, p\rangle \), the resulting mixture would be \([\text{Stapp (1995)}]\)

\[
\int \mu(q, p) |q, p\rangle \langle q, p| \langle q, p|\Psi\rangle \langle \Psi|q, p\rangle \langle q, p|,
\]

whose trace is

\[
\int \mu(q, p) \langle q, p|\Psi\rangle \langle \Psi|q, p\rangle = \langle \Psi|\Psi\rangle.
\]

To represent the limited capacity of consciousness let us assume, in this model, that the states of consciousness associated with a brain can be expressed in terms of a relatively small subset of the modes of the electromagnetic field in the brain cavity. Let us assume that events occurring outside the brain are keeping the state of the universe outside the brain cavity in a single state, so that the state of the brain can also be represented by a single state. The brain is represented, in the path–integral method of Feynman, by a superposition of the trajectories of the particles in it, with each element of the superposition accompanied by the coherent–state electromagnetic field that this set of trajectories generates. Let the state of the electromagnetic field restricted to the modes that represent consciousness be called \( |\Psi(t)\rangle \). Using the decomposition of unity one can write

\[
|\Psi(t)\rangle = \int \mu(q, p) |q, p\rangle \langle q, p|\Psi(t)\rangle.
\]

Hence the state at time \( t \) can be represented by the function \( \langle q, p|\Psi(t)\rangle \), which is a complex-valued function over the set of arguments \( \{q_1, p_1, q_2, p_2, ..., q_n, p_n\} \), where \( n \) is the number of modes associated with \( |\Psi\rangle \). Thus in this model the contents of the consciousness associated with a brain is represented in terms of this function defined over a 2nD space: the \( i \)th conscious event is represented by the transition

\[
|\Psi_i(t_{i+1})\rangle \rightarrow |\Psi_{i+1}(t_{i+1})\rangle = P_i|\Psi_i(t_{i+1})\rangle,
\]

where \( P_i \) is a projection operator.
For each allowed value of $k$ the pair of numbers $(q_k, p_k)$ represents the state of motion of the $k$th mode of the electromagnetic field. Each of these modes is defined by a particular wave pattern that extends over the whole brain cavity. This pattern is an oscillating structure something like a sine wave or a cosine wave. Each mode is fed by the motions of all of the charged particles in the brain. Thus each mode is a representation of a certain integrated aspect of the activity of the brain, and the collection of values $q_1, p_1, \ldots, p_n$ is a compact representation of certain aspects the over-all activity of the brain.

The state $|q, p\rangle$ represents the conjunction, or collection over the set of all allowed values of $k$, of the various states $|q_k, p_k\rangle$. The function

$$V(q, p, t) = \langle q, p | \Psi(t) \rangle \langle \Psi(t) | q, p \rangle$$

satisfies $0 \leq V(q, p, t) \leq 1$, and it represents, according to orthodox thinking, the ‘probability’ that a system that is represented by a general state $|\Psi(t)\rangle$ just before the time $t$ will be observed to be in the classically describable state $|q, p\rangle$ if the observation occurs at time $t$. The coherent states $|q, p\rangle$ can, for various mathematical and physical reasons, be regarded as the ‘most classical’ of the possible states of the electromagnetic quantum field.

To formulate a causal dynamics in which the state of consciousness itself controls the selection of the next state of consciousness one must specify a rule that determines, in terms of the evolving state $|\Psi(t)\rangle$ up to time $t_{i+1}$, both the time $t_{i+1}$ when the next selection event occurs, and the state $|\Psi_{i+1}(t_{i+1})\rangle$ that is selected and actualized by that event.

In the absence of interactions, and under certain ideal conditions of confinement, the deterministic normal law of evolution entails that in each mode $k$ there is an independent rotation in the $(q_k, p_k)$ plane with a characteristic angular velocity $\omega_k = k_0$. Due to the effects of the motions of the particles there will be, added to this, a flow of probability that will tend to concentrate the probability in the neighborhoods of a certain set of ‘optimal’ classical states $|q, p\rangle$. The reason is that the function of brain dynamics is to produce some single template for action, and to be effective this template must be a ‘classical’ state, because, according to orthodox ideas, only these can be dynamically robust in the room temperature brain. According to the semi-classical description of the brain dynamics, only one of these classical-type states will be present, but according to quantum theory there must be a superposition of many such classical-type states, unless collapses
Quantum Leap

occurs at lower (i.e., microscopic) levels. The assumption here is that no
collapses occur at the lower brain levels: there is absolutely no empirical
evidence, or theoretical reason, for the occurrence of such lower–level brain
events.

So in this model the probability will begin to concentrate around various
locally optimal coherent states, and hence around the various (generally)
isolated points \((q, p)\) in the 2nD space at which the quantity \[ V(q, p, t) = \langle q, p | \Psi(t) \rangle \langle \Psi(t) | q, p \rangle \]
reaches a local maximum. Each of these points \((q, p)\) represents a locally–
optimal solution (at time \(t\)) to the search problem: as far as the myopic local
mechanical process can see the state \(|q, p\rangle\) specifies an analog-computed
‘best’ template for action in the circumstances in which the organism finds
itself. This action can be either intentional (it tends to create in the future
a certain state of the body/brain/environment complex) or attentional (it
tends to gather information), and the latter action is a special case of the
former. As discussed in \[Stapp (1993)\], the intentional and attentional
character of these actions is a consequence of the fact that the template for
action actualized by the quantum brain event is represented as a projected
body–world schema, i.e., as the brains projected representation of the body
that it is controlling and the environment in which it is situated.

Let a certain time \(t_{i+1}|t_i\) be defined by an (urgency) energy factor
\(E(t) = h(t_{i+1} - t_i)^{-1}\). Let the value of \((q, p)\) at the largest of the local–
maxima of \(V(q, p, t_{i+1})\) be called \((q(t_{i+1}), p(t_{i+1}))_{\text{max}}\). Then the simplest
possible reasonable selection rule would be given by the formula
\[ P_i = |(q(t_{i+1}), p(t_{i+1}))_{\text{max}}\rangle \langle (q(t_{i+1}), p(t_{i+1}))_{\text{max}}|, \]
which entails that
\[ |\Psi_{i+1}\rangle\langle \Psi_{i+1} | = |(q(t_{i+1}), p(t_{i+1}))_{\text{max}}\rangle \langle (q(t_{i+1}), p(t_{i+1}))_{\text{max}}|. \]

This rule could produce a tremendous speed up of the search process.
Instead of waiting until all the probability gets concentrated in one state
\(|q, p\rangle\), or into a set of isolated states \(|q_i, p_i\rangle\) [or choosing the state randomly,
in accordance with the probability function \(V(q, p, t_{i+1})\), which could often
lead to a disastrous result], this simplest selection process would pick the
state \(|q, p\rangle\) with the largest value of \(V(q, p, t)\) at the time \(t = t_{i+1}\). This
process does not involve the complex notion of picking a random number,
which is a physically impossible feat that is difficult even to define.
One important feature of this selection process is that it involves the state $\Psi(t)$ as a whole: the whole function $V(q, p, t_{i+1})$ must be known in order to determine where its maximum lies. This kind of selection process is not available in the semi-classical ontology, in which only one classically describable state exists at the macroscopic level. That is because this single classically describable macro-state state (e.g., some one actual state $|q, p, t_{i+1}\rangle$) contains no information about what the probabilities associated either with itself or with the other alternative possibilities would have been if the collapse had not occurred earlier, at some micro-level, and reduced the earlier state to some single classically describable state, in which, for example, the action potential along each nerve is specified by a well defined classically describable electromagnetic field. There is no rational reason in quantum mechanics for such a micro-level event to occur. Indeed, the only reason to postulate the occurrence of such premature reductions is to assuage the classical intuition that the action-potential pulse along each nerve 'ought to be classically describable even when it is not observed', instead of being controlled, when unobserved, by the local deterministic equations of quantum field theory. But the validity of this classical intuition is questionable if it severely curtails the ability of the brain to function optimally.

A second important feature of this selection process is that the actualized state $\Psi_{i+1}$ is the state of the entire aspect of the brain that is connected to consciousness. So the feel of the conscious event will involve that aspect of the brain, taken as a whole. The ‘I’ part of the state $\Psi(t)$ is its slowly changing part. This part is being continually re-actualized by the sequence of events, and hence specifies the slowly changing background part of the felt experience. It is this persisting stable background part of the sequence of templates for action that is providing the over-all guidance for the entire sequence of selection events that is controlling the on-going brain process itself [Stapp (1995)].

A somewhat more sophisticated search procedure would be to find the state $|q, p\rangle_{max}$, as before, but to identify it as merely a candidate that is to be examined for its concordance with the objectives imbedded in the current template. This is what a good search procedure ought to do: first pick out the top candidate by means of a mechanical process, but then evaluate this candidate by a more refined procedure that could block its acceptance if it does not meet specified criteria.

It may at first seem strange to imagine that nature could operate in such a sophisticated way. But it must be remembered that the generation of a
truly random sequence is itself a very sophisticated (and indeed physically impossible) process, and that what the physical sciences have understood, so far, is only the mechanical part of nature’s two-part process. Here it is the not-well-understood selection process that is under consideration. We have imposed on this attempt to understand the selection process the naturalistic requirement that the whole process be expressible in natural terms, i.e., that the universal process be a causal self-controlling evolution of the Hilbert-space state-vector in which all aspects of nature, including our conscious experiences, are efficacious.

It may be useful to describe the main features of this model in simple terms. If we imagine the brain to be, for example, a uniform rectangular box then each mode \( k \) would correspond to wave form that is periodic in all three directions: it would be formed as a combination of products of sine waves and cosine waves, and would cover the whole box-shaped brain. (More realistic conditions are needed, but this is a simple prototype.) Classically there would be an amplitude for this wave, and in the absence of interactions with the charged particles this amplitude would undergo a simple periodic motion in time. In analogy with the coordinate and momentum variables of an oscillating pendulum there are two variables, \( q_k \) and \( p_k \), that describe the motion of the amplitude of the mode \( k \). With a proper choice of scales for the variables \( q_k \) and \( p_k \) the motion of the amplitude of mode \( k \) if it were not coupled to the charges would be a circular motion in the \((q_k, p_k)\) plane. The classical theory would say that the physical system, mode \( k \), would be represented by a point in \( q_k, p_k \) space. But quantum theory says that the physical system, mode \( k \), must be represented by a wave (i.e., by a wave \( \psi \)-function) in \((q_k, p_k)\) space. The reason is that interference effects between the values of this wave (function) at different points \((q_k, p_k)\) can be exhibited, and therefore it is not possible to say the full reality is represented by any single value of \((q_k, p_k)\): one must acknowledge the reality of the whole wave. It is possible to associate something like a ‘probability density’ with this wave, but the corresponding probability cannot be concentrated at a point: in units where Planck’s constant is unity the bulk of the probability cannot be squeezed into a region of the \((q_k, p_k)\) plane of area less that unity.

The mode \( k \) has certain natural states called ‘coherent states’, \(|q_k, p_k\rangle\). Each of these is represented in \((q_k, p_k)\)-space by a wave function that has a ‘probability density’ that falls off exponentially as one moves in any direction away from the center-point \((q_k, p_k)\) at which the probability density is maximum. These coherent states are in many ways the ‘most
classical’ wave functions allowed by quantum theory \cite{Glauber1963a}, \cite{Glauber1963b}, and a central idea of the present model is to specify that it is to one of these ‘most classical’ states that the mode-$k$ component of the electromagnetic field will jump, or collapse, when an observation occurs. This specification represents a certain ‘maximal’ principle: the second process, which is supposed to pick out and actualize some classically describable reality, is required to pick out and actualize one of these ‘most classical’ of the quantum states. If this selection/actualization process really exists in nature then the classically describable states that are actualized by this process should be ‘natural classical states’ from some point of view. The coherent states satisfy this requirement. This strong, specific postulate should be easier to disprove, if it is incorrect, than a vague or loosely defined one.

If we consider a system consisting of a collection of modes $k$, then the generalization of the single coherent state $|q_k,p_k\rangle$ is the product of these states, $|q,p\rangle$. Classically this system would be described by specifying the values all of the classical variables $q_k$ and $p_k$ as functions of time. But the ‘best’ that can be done quantum mechanically is to specify that at certain times $t_i$ the system is in one of the coherent states $|q,p\rangle$. However, the equations of local quantum field theory (here quantum electrodynamics) entail that if the system starts in such a state then the system will, if no ‘observation’ occurs, soon evolve into a superposition (i.e., a linear combination) of many such states. But the next ‘observation’ will then reduce it again to some classically describable state. In the present model each a human observation is identified as a human conscious experience. Indeed, these are the same observations that the pragmatic Copenhagen interpretation of Bohr refers to, basically. The ‘happening’ in a human brain that corresponds to such an observation is, according to the present model, the selection and actualization of the corresponding coherent state $|q,p\rangle$.

The quantity $V(q,p,t_{i+1})$ defined above is, according to orthodox quantum theory, the predicted probability that a system that is in the state $\Psi(t_{i+1})$ at time $t_{i+1}$ will be observed to be in state $|q,p\rangle$ if the observation occurs at time $t_{i+1}$. In the present model the function $V(q,p,t_{i+1})$ is used to specify not a fundamentally stochastic (i.e., random or chance–controlled) process but rather the causal process of the selection and actualization of some particular state $|q,p\rangle$. And this causal process is controlled by features of the quantum brain that are specified by the Hilbert space representation of the conscious process itself. This process is a nonlocal process that rides
on the local brain process, and it is the nonlocal selection process that, according to the principles of quantum theory, is required to enter whenever an observation occurs.

7.1.6 Theory of Mind and Matter

Most of the physicists today believe that we are very close to a unified theory of matter (UTM) based on $p$-brane theory, a generalization of superstring theory (see, e.g., Ivancevic and Ivancevic (2006b)). This theory has been well motivated by the successes (from microchips to satellites) and the difficulties (the presence of infinities) of Quantum Gauge Field Theories (QED, Electro–Weak theory, GUT and SUSY GUT etc.) most of which are based on solid experimental evidences (measuring Lande $g$–factor of electron up to 10 significant places, prediction of $Z$ boson etc.). This ‘theory of everything’ tells us that all carbon atoms (after they are born out of nucleosynthesis in the core of stars) in this universe are the same in space and in time, and thus is unable to explain why the carbon atoms present in a lump of roughly three pound of ordinary matter the human brain – give rise to such ineffable qualities as feeling, thought, purpose, awareness and free will that are taken to be evidence for ‘consciousness’.

Is Consciousness an accident caused by random evolutionary processes in the sense that it would not evolve again had the present universe to undergo a ‘Big Crunch’ to start with another ‘Big Bang’? No. It is a fundamental property that emerges as a natural consequence of laws of nature. Are these laws of nature different from laws of physics? Is there a way to expand the UTM to incorporate consciousness?

Nature manifests itself not only at the gross level of phenomena (accessible to direct senses) but also at the subtle level of natural laws (accessible to ‘refined’ senses). Consciousness is the ability to access nature at both these levels. Hence everything in nature is conscious, but there is a hierarchy in the level of consciousness. Although animals, plants and few machines today can access to the gross level of nature, access to the subtle level seems to be purely human. In this sense a layman is less conscious compared to a scientist or an artist. (Is it possible that a level of consciousness exists compared to which a scientist or an artist of today may appear a layman?) Once everything (both matter and mind) is reduced to information it is possible to define consciousness as the capability to process information. Because any form of our access to nature is based on processing information with varying degree of complexity. The words process and complexity will be defined in the following subsection, mainly following Samal (2001).
7.1.6.1 Matter

Phenomena

Some animals like dogs (in listening to ultrasound) and bats (in sensing prey through echo technique) are better equipped than humans when it comes to direct sense experiences. But unlike animals, humans have found out methods to extend their sense experiences through amplification devices like telescopes, microscopes etc. The summary of such sense experiences (direct and extended) acquired over the last five hundred years (equivalent to one second in a time-scale where the age of the universe is equal to a year) is that our universe extends in space from the size of an electron \(10^{17} \text{ cm}\) as probed at LEP colliders at CERN, Geneva to size of galactic super-clusters \(10^{30} \text{ cm}\) as probed by high red-shift measurements. The theoretical possibility of Planck length \(10^{33} \text{ cm}\) allows for further extension in space.

Our grand universe extending over more than 60 orders of magnitude in space has been constantly changing over the last 12 billion years. Living systems seem to have evolved out of non-living systems. Consciousness seems to have evolved out of living systems. Is the dramatic difference between animate and inanimate, conscious and unconscious simply a difference in their ability to process information? At the level of phenomena nature is so vast and diverse compared to the size and comprehension of human beings that one wonders how the collective inquiring human minds over the centuries could at all fathom the unity behind this diversity. This simply testifies the triumph of human mind over the physical limitations of its body. Because the human mind is capable of reaching a synthesis (an advanced form of information processing) based on careful observations of diverse phenomena.

Inanimate Manifold

Till Einstein, everybody thought ‘absolute space’ is the arena (mathematically, a manifold) on which things change with ‘absolute time’. In special theory of relativity (STR) he redefined the manifold to be flat space-time \((3+1)\) by making both space and time relative with respect to inertial observers but keeping space-time \((\text{SpT})\) absolute. In general theory of relativity (GTR) he propounded this manifold to be curved space-time that
can act on matter unlike the flat space-time. Then the Quantum Theory (QT, the standard version, not the Bohmian one) pointed out this manifold to be an abstract mathematical space called Hilbert space since the space-time description of quantum processes is not available. Quantum Field Theory (QFT) that originated in a successful merge of QT and STR requires this manifold to be the Quantum Vacuum (QV). Unlike the ordinary vacuum, QV contains infinite number of ‘virtual’ particles that give rise to all matter and interactions. Every quantum field has its own QV and a successful amalgamation of QT and GTR (called Quantum Gravity, and yet to be achieved) may connect the curved space-time with the QV of gravitational field. Finally, Superstring (or, p–brane) theory demands all fundamental entities to be strings (or p–branes) in 10D space-time. This evolution in our understanding clearly shows the necessity and importance of a background manifold to formulate any scientific theory.

Basic Constituents

Energy and matter were considered to be the two basic constituents of our universe till Einstein showed their equivalence through $E = mc^2$. All forms of energy are interconvertible. Matter in all its forms (solid, liquid, gas and plasma) consists of atoms. The simplest of all atoms is the hydrogen atom that contains an electron and a proton. If one considers the (now outdated) Bohrian picture of hydrogen atom being a miniature solar system where the electron revolves around the proton in circular orbit then one realizes that most of the hydrogen atom is empty space. This means that if one blows up the hydrogen atom in imagination such that both electron and proton acquire the size of a football each then they need to be separated by 100 km or more. Hence one would think that even if 99.99% of the hydrogen atom consists of vacuum the point–like electron and proton are material particles. But that is not so.

The elementary particle physics tells us that all the visible matter (composition of dark matter is not yet surely known) in the universe is made of six leptons and six quarks along with their antiparticles. Although they are loosely called particles they are not like the ordinary particles we experience in daily life. A ‘classical’ particle is localized and impenetrable whereas a ‘classical’ wave can be extended from minus infinity to plus infinity in space and many wave modes can simultaneously occupy the same space (like inside the telephone cable). But with the advent of quantum mechanics this seemingly contradictory differences between particle and wave lost their sharpness in the quantum world. A quantum object can simultaneously ‘be’ a particle and a wave until a measurement is made on it. According
to QFT all fundamental entities are quantum fields (not material in the conventional sense) that are neither particle nor wave in the classical sense.

**Evolution**

Despite various specializations, all physical sciences share a common goal: given the complete specification of a physical system at an initial time (called the initial conditions) how to predict what will it be at a later time. To predict with exactness one needs the accurate initial conditions and the laws that govern the evolution of the system (called the dynamical laws). Either the lack of exact specifications of initial conditions or the intractability of huge number of equations (that express dynamical laws) can lead to a probabilistic description rather than a deterministic one. However, there are chaotic systems that can be both deterministic yet unpredictable because of their extreme sensitivity to initial conditions. Apart from dynamical laws there are other laws like Einstein’s \( E = mc^2 \), etc., which do not involve time explicitly. Could there be laws at the level of initial conditions that guide us to choose a particular set over another?

A physical law is like the hidden thread (unity) of a garland with various flowers representing diverse natural phenomena. Its character seems to depend on characteristic scales (denoted by fundamental constants like Planck length, Planck’s constant, and speed of light etc). Why should there be different set of laws at different scales? Most physicists believe that quantum mechanics is universal in the applicability of its laws like Schrödinger’s equation and classicality of the everyday world is a limiting case. But there exists no consensus at present regarding the emergence of this limiting case.

**Guiding Principles**

How does one formulate these physical laws? The principle of relativistic causality helps. Do physical laws change? Is there a unity behind the diversity of laws? Is it possible to understand nature without laws? The concept of symmetry (invariance) with its rigorous mathematical formulation and generalization has guided us to know the most fundamental of physical laws. Symmetry as a concept has helped mankind not only to define ‘beauty’ but also to express the ‘truth’. Physical laws tries to quantify the truth that appears to be ‘transient’ at the level of phenomena but symmetry promotes that truth to the level of ‘eternity’.

**Interactions**

The myriad mosaic of natural phenomena is possible because, not only each fundamental entity evolves with time but also it can interact with the other basic constituents. All the physical interactions (known so far) can
be put into four categories:
(i) gravitational interaction (the force that holds the universe),
(ii) electromagnetic interaction (the force that holds the atom, and hence all of us),
(iii) strong nuclear interaction (the force that holds the nucleus), and
(iv) weak nuclear interaction (the force that causes radioactive decay).

At present (ii) and (iv) are known (observationally) to be unified to a single force called Electro–Weak. Grand unified theories (GUT) and their extensions (SUSY GUT) for (ii), (iii) and (iv) do exist. Ongoing research aims to unify (i) with such theories.

According to QFT the basic matter fields interact by exchanging messenger fields (technically called gauge fields) that define the most fundamental level of communication in nature. Both matter fields and gauge fields originate in the fluctuations of QV and in this sense everything in universe including consciousness is, in principle, reducible to QV and its fluctuations. Communication in nature can happen either via the local channel mediated by gauge fields or by the nonlocal EPR [Einstein et al. (1935a)] type channels (through entanglement) as was demonstrated by recent quantum teleportation experiments.

Composite Systems
Till the importance of quantum entanglement was realized in recent times the whole was believed to be just the sum of parts. But the whole seems to be much more than just the sum of parts in the ‘quantum’ world as well as classical systems having complexity. It makes quantum entanglement a very powerful resource that has been utilized in recent times for practical schemes like quantum teleportation, quantum cryptography and quantum computation. Quantum nonlocality indicates that the universe may very well be holographic in the sense that the whole is reflected in each part [Samal (2001)].

Animate

Manifold
A (3 + 1)—space-time is the manifold for all biological functions at the phenomenal level that can be explained by classical physics. If one aims to have a quantum physical explanation of certain biological functions then the manifold has to be the Hilbert space.

Basic Constituent
Cell is the basic constituent of life although the relevant information
seems to be coded at the subcell (genetic) level. Neuron (or, microtubules and cytoskeletons) could be the physical substratum of brain depending on classical (or, quantum) viewpoint.

Evolution

Does biological evolution happen with respect to the physical time? If yes, then will the physical laws suffice to study biological evolution in the sense they do in chemistry? If no, then is the biological arrow of time different from the various arrows of physical time (say, cosmological or the thermodynamical arrow of time)? Is there a need for biological laws apart from physical laws to understand the functioning of biological systems?

Guiding Principles

Survivability is the guiding principles in biological systems. Organisms constantly adapt to each other through evolution, and thus organizing themselves into a delicately tuned ecosystem. Intentionality may also play a very important role in the case of more complex bio–systems.

Interactions

The interaction occurs by exchange of chemicals, electric signals, gestures, and language etc. at various levels depending upon the level of complexity involved.

Composite Systems

Composite systems are built out of the basic constituents retaining the relevant information in a holographic manner. The genetic information in the zygote is believed to contain all the details of the biology to come up later when the person grows up. The genes in a developing embryo organize themselves in one way to make a liver cell and in another way to make a muscle cell [Samal (2001)].

Discussions

How ‘material’ is physical? Anything that is physical need not be ‘material’ in the sense we experience material things in everyday life. The concept of energy is physical but not material. Because nobody can experience energy directly, one can only experience the manifestations of energy through matter. Similarly the concept of a ‘classical field’ in physics is very abstract and can only be understood in terms of analogies. Still more abstract is the concept of a ‘quantum field’ because it cannot be understood in terms of any classical analogies. But at the same time it is a well–known fact in modern physics that all fundamental entities in the universe are quantum fields. Hence one has to abandon the prejudice that anything ‘physical’ has
to be ‘material’.

Is reductionism enough? The reductionist approach: observing a system with an increased resolution in search of its basic constituents has helped modern science to be tremendously successful. The success of modern science is the success of the experimental method that has reached an extreme accuracy and reproducibility. But the inadequacy of reductionism in physical sciences becomes apparent in two cases: emergent phenomena and quantum nonlocality. Quantum nonlocality implies a holographic universe that necessitates a holistic approach [Bohm and Hiley (1993)].

Though it is gratifying to discover that everything can be traced back to a small number of quantum fields and dynamical laws it does not mean that we now understand the origin of earthquakes, weather variations, the growing of trees, the fluctuations of stock market, the population growth and the evolution of life? Because each of these processes refers to a system that is complex, in the sense that a great many independent agents are interacting with each other in a great many ways. These complex systems are adaptive and undergo spontaneous self–organization (essentially nonlinear) that makes them dynamic in a qualitatively different sense from static objects such as computer chips or snowflakes, which are merely complicated. Complexity deals with emergent phenomena. The concept of complexity is closely related to that of understanding, in so far as the latter is based upon the accuracy of model descriptions of the system obtained using condensed information about it [Badii and Politi (1997)].

In this sense there are three ultimate frontiers of modern physics: the very small, the very large and the very complex. Complex systems cease to be merely complicated when they display coherent behavior involving collective organization of vast number of degrees–of–freedom. Wetness of water is a collective phenomenon because individual water molecules cannot be said to possess wetness. Lasers, superfluidity and superconductivity are few of the spectacular examples of complexity in macroscopic systems, which cannot be understood alone in terms of the microscopic constituents. In every case, groups of entities seeking mutual accommodation and self–organization somehow manage to transcend the individuality in the sense that they acquire collective properties that they might never have possessed individually. In contrast to the linear, reductionist thinking, complexity involves nonlinearity and chaos and we are at present far from understanding the complexity in inanimate processes let alone the complexity in living systems.
Emergence of Life

Is life nothing more than a particularly complicated kind of carbon chemistry? Or is it something subtler than putting together the chemical components? Do computer viruses have life in some fundamental sense or are they just pesky imitations of life? How does life emerge from the quadrillions of chemically reacting proteins, lipids, and nucleic acids that make up a living cell? Is it similar to the emergence of thought out of the billions of interconnected neurons that make up the brain? One hope to find the answer to these questions once the dynamics of complexity in inanimate systems is well understood [Samal (2001)].

7.1.6.2 Mind

Phenomena

Mind is having three states: awake, dream, dreamless sleep. Mind is capable of free-will, self-perception (reflective) and universal perception (perceptual) in its ‘awake’ state. Can it be trained to have all these three attributes in the states of dream and dreamless sleep? Can there be a fourth state of mind that transcends all the above three states? Where do the brain end and the mind begin? Due to its global nature, mind cannot lie in any particular portion of the brain. Does it lie everywhere in the brain? This would require nonlocal interactions among various components of the brain. If there were no such nonlocal communication then how does the mind emerge from the brain?

Can anybody think of anything that transcends space-time? Is mind capable of thinking something absolutely new that has not been experienced (directly or indirectly) by the body? Nobody can think of anything absolutely new. One can only think of a new way of arranging and/or connecting things that one has ever learnt. In this sense intellect is constrained by reason whereas imagination is not. But imagination is not acceptable to intellect unless it is logically consistent with what is already known. Imagination helps to see a new connection but intellect makes sure that the new connection is consistent with the old structure of knowledge. This is the way a new structure in knowledge is born and this process of acquiring larger and larger structure (hence meaning or synthesis) is the learning process. Science is considered so reliable because it has a stringent methodology to check this consistency of imagination with old knowledge.

Can one aspire to study the mind using methodology of (physical) sci-
ences? Seeing the tremendous success of physical sciences in the external world one would think its methodology to work for understanding the inner world. It is not obvious \textit{a priori} why should not QT work in this third ontology when it has worked so successfully with two different ontologies? We aim to understand nature at a level that transcends the inner and the outer worlds by synthesizing them into a more fundamental world of quantum information \cite{Samal(2001)}.

\section*{Formalism}

\textit{Manifold}

A physical space-time description of mind is not possible because thoughts that constitute the mind are acausal: it does not take 8 minutes for me to think of the sun although when I look at the sun I see how it was 8 minutes ago. We will assume that it is possible to define an abstract manifold for the space of thoughts (say, $T$-space). An element of $T$-space is a thought-state ($T$-state) and the manifold allows for a continuous change from one $T$-state to another. I presume that $T$-space is identical with mind but it need not be so if mind can exist in a thoughtless but awake state, called turiya state.

\textit{Basic Constituent}

How does one define a $T$-state? That requires one to understand what is a ‘thought’? A thought always begins as an idea (that could be based on self and universal perception) and then undergoes successive changes in that idea but roughly remaining focused on a theme. Change from one theme to another is triggered by a new idea. Hence I would suggest that the basic constituent of $T$-state is idea. An idea is like a ‘snapshot’ of experience complete with all sense data whereas a thought ($T$-state) is like an ensemble (where each element is not an exact replica of the other but has to be very close copy to retain the focus on the theme) of such snapshots.

\textit{Evolution}

There are two types of evolution in $T$-space. First, the way an ensemble of ideas evolves retaining a common theme to produce a thought. To concentrate means to linger the focus on that theme. This evolution seems to be nonlinear and nondeterministic. Hence the linear unitary evolution of QT may not suffice to quantify this and will be perhaps best described in terms of the mathematics of self-organization and far-from-
equilibrium phenomena. The second type involves a change from one particular thought into another and this evolution could be linear and perhaps can be calculated through a probability amplitude description in the line of QT. Given the complete description of a thought at an initial time the refutability of any theory of mind amounts to checking how correctly it can predict the evolution of that thought at a later time.

**Guiding Principles**

If the guiding principle for evolution in biological world is survivability then in $T^{-}$–space it is happiness–ability. A constant pursuit of happiness (although its definition may vary from person to person) guides the change in a person’s thoughts. Each and every activity (begins as mental but may or may not materialize) is directed to procure more and more happiness in terms of sensual pleasures of the body, emotional joys of the imagination and rational delights of the intellect.

**Interactions**

Can a thought (mind) interact with another thought (mind)? Can this interaction be similar to that between quantum fields? Perhaps yes, only if both thought and quantum fields can be reduced to the same basic entity. Then it will be possible for thought (mind) to interact with matter. What will be the messenger that has to be exchanged between interacting minds or between interacting mind and matter? This ultimate level of communication has to be at the level of QV and hence it may amount to silence in terms of conventional languages. But can any receiver (either human mind or any other mind or equipment) be made so sensitive to work with this ultimate level of communication? Interaction with the environment is believed to decohere a quantum system that causes the emergence of classicality in physical world. A completely isolated system remains quantum mechanical. Can a completely isolated mind exhibit quantum mechanical behavior in the sense of superposition and entanglement?

**Composite Systems**

In the $T^{-}$–Space a thought is an ensemble of ideas and a mind–state is composed of thoughts. Behavior, feeling and knowledge of self and universe are in principle reducible to composite subsets in the $T^{-}$–space.  

---

[33] On the practical side the time–tested techniques of Yoga teach us how to linger the focus on a theme through the practice of *concentration*, *meditation* and *samadhi*.  

---
Discussions

*Working Definition of Consciousness*

Consciousness (at the first level) is related to one’s response $R$ to one’s environment. This response consists of two parts: habit $H$ and learning $L$. Once something is learnt and becomes a habit it seems to drop out of consciousness. Once driving a bicycle is learnt one can think of something else while riding the bicycle. But if the habit changes with time then it requires conscious attention. We have defined learning earlier as a process to find commensurability of a new experience with old knowledge. One has to learn anew each time there is a change in the environment. Hence consciousness is not the response to the environment but is the time of rate of change of the response $C = \partial_t R$, where $R = H + L$.

The hierarchy in consciousness depends on the magnitude of this time derivative. Everything in the universe can be fit in a scale of consciousness with unconscious and super-conscious as the limit points. It is obvious that all animals show response to their environment, so does some of the refrigerators, but there is a hierarchy in their response. Through the use of ‘cresco-graph’ and ‘resonant cardio-graph’ of J.C. Bose, one can see the response of botanical as well as inanimate world. We cannot conclude that a stone is unconscious just because we cannot communicate with it using our known means of communication. As technology progresses, we will be able to measure both the response function $R$ and its time derivative. If this is the definition what can it tell about the future evolution of humans? My guess is that we would evolve from conscious to super-conscious in the sense that genes will evolve to store the cumulative learning of the human race.

*Emergence of Consciousness*

The first step in understanding consciousness consists of using reductionist method to various attributes of consciousness. A major part of the studies done by psychologists (and their equivalents doing studies on animals) and neurobiologists falls under this category. Such studies can provide knowledge about mind states (say, $M_1, M_2, M_3, ...$) but cannot explain the connection between these mind states with the corresponding brain states (say $B_1, B_2, ...$). Because this kind of dualistic model of Descartes would require to answer a) where is mind located in the brain, and b) if my mind wants me to raise my finger, how does it manage to trigger the appropriate
nerves and so on in order for that to happen without exerting any known forces of nature?

To find out how the mind actually works one needs to have a theory of mind, that will relate the sequence of mental states $M_1, M_2, M_3, \ldots$ by providing laws of change (the dynamical laws for the two types of evolution discussed above) that encompass the mental realm after the fashion of the theory of matter that applies to the physical realm, with its specific laws. Such a theory of mind is possible if we synthesize the results of studies on attributes of consciousness to define the exact nature of the manifold and the basic entities of the $T-$space (or, Mind–Space). Once this is achieved then one can attempt to explain the emergence of consciousness taking clues from complexity theory in physical sciences. But such an extrapolation will make sense provided both $M-$states and $B-$states can be reduced to something fundamental that obeys laws of complexity theory. We propose in the next section that information is the right candidate for such a reduction.

**Role of Indian Philosophy (IP)**

(1) Unlike the Cartesian dichotomy of mind and body some schools of IP like Vaisheshika and Yoga treat both mind and body in a unified manner. Since (western) science is based on Cartesian paradigm it cannot synthesize mind and body unless it takes the clue from oriental philosophies and then blend it with its own rigorous methodology.

(2) In terms of sense awareness, awake, dream and dream–less sleep states are often called as conscious, subconscious and unconscious states. A great conceptual step taken by IP in this regard is to introduce a fourth state of mind called *turiya state* that is defined to be none of the above but a combination of all of the above states. This state is claimed to be the super–conscious state where one transcends the limitations of perceptions constrained by space-time $(3 + 1)$. Patanjali has provided very scientific and step by step instructions to reach this fourth state through samyama (concentration, meditation and samadhi are different levels of samyama). The scientific validity of this prescription can be easily checked by controlled experiments. Nobody can understand the modern physics without going through the prerequisite mathematical training. It will be foolish for any intelligent lay person to doubt the truth of modern physics without first undergoing the necessary training. Similarly one should draw conclusion about yogic methods only after disciplined practice of the eight steps of yoga.

(3) IP can provide insights regarding the role of mind in getting happiness and thus a better understanding of mind itself. Happiness lies in what the
mind perceives as pleasurable and hence the true essence of happiness lies in mind and not in any external things. Once the body has experienced something mind is capable of recreating that experience in the absence of the actual conditions that gave rise to the experience in the first place. One can use this capacity of mind to create misery or ecstasy depending on one’s ability to guide one’s mind.

(4) There is a concept of the primordial sound in IP. Sometimes the possibility of having a universal language to communicate with everything in the universe is also mentioned. Modern physics tells us that the only universal language is at the level of gauge bosons and QV. Is there any connection between these two? Can a human mind be trained to transmit and receive at the level of QV?

(5) It is said that whole body is in the mind whereas the whole mind is not in the body. How does mind affect the body? If one believes in the answer given by IP then the results obtained in this regard by the western psychology appears to be the tip of the iceberg only. Can science verify these oriental claims through stringently controlled experiments?

7.1.6.3 Unification

Information

Information seems to be abstract and not real in the sense that, it lies inside our heads. But information can, not only exist outside the human brain (i.e., library, a CD, internet etc.) but also can be processed outside human brain (i.e., other animals, computers, etc.). Imagine a book written in a dead language, which nobody today can decipher. Does it contain information? Yes. Information exists. It does not need to be perceived or understood to exist. It requires no intelligence to interpret it. In this sense information is as real as matter and energy when it comes to the internal structure of the universe [Stonier (1990)]. But what we assume here is that information is more fundamental than matter and energy because everything in the universe can be ultimately reduced to information.

Information is neither material nor non–material. Both, quantum fields and thoughts can be reduced to information. If the human mind is not capable (by the methods known at present) of understanding this ultimate information then it is the limitation of the human mind. This may not remain so as time progresses. The whole of physical world can be reduced to information [Frieden (1999)]. Is information classical or quantum? There
are enough indications from modern physics that although it can be classical at the everyday world it is quantum at the most fundamental level. The quantum information may have the advantage of describing the fuzziness of our experiences.

**Formalism**

*Mannifold*

The manifold is an information field (I-field) for classical information (like that of Shannon or Fisher, etc.) Hilbert space of QT is the manifold to study quantum information. But if quantum information has to be given an ontological reality then it may be necessary for the manifold to be an extended Hilbert space.

*Basic Constituent*

A bit or a qubit is the basic entity of information depending on whether it is treated as classical or quantum respectively. Information can be of two types: kinetic and structural, but they are convertible to each other [Stonier (1990)].

*Evolution*

All organized systems contain information and addition of information to a system manifests itself by causing the system to become more organized or reorganized. The laws for evolution of information are essentially laws of organization. Are these laws different from the physical laws? Is there an equivalent in the world of information of fundamental principles like principle of least action in physical world?

*Guiding Principles*

Optimization seems to be the guiding principle in the world of information. What gives rise to the structure in the information such that we acquire an understanding or meaning out of it? Is there a principle of least information to be satisfied by all feasible structures?

*Interactions*

The interaction at the level of information has to be the ultimate universal language. What could be that language? The only fundamental language known to us is that of the gauge fields that communicate at the level of QV. Could the gauge fields serve as quanta of information? How far is this language from the conventional language? Can this help us to communicate with not only with other creatures incapable of our conventional language but also with the inanimate world? Time is not yet ripe to answer these questions.
Composite Systems
How can every composite system of information (like a gene, or a galaxy) be expressed in terms of bits or qubits? Does the holographic principle also apply to information?

Discussions

Consciousness and Information
There is no doubt that sooner or later all attributes of consciousness can be reduced to information. This is just a matter of time and progress in technology. That will complete the understanding of consciousness at the gross level of phenomena but will harbinger the understanding of consciousness at the subtle level of laws. The synthesis of the phenomenological studies of consciousness will be possible by treating information as the most basic ontological entity, which can unify mind and matter. The emergence of consciousness will be understood in terms of nonlinear, far-from-equilibrium complex processes that lead to spontaneous self-organization and adaptation of structures in the manifold of quantum information.

Consciousness will be seen as the ability to process quantum information in an effective way. Depending on the degree of complexity involved the processing would encompass activities starting from the way a planet knows which is the path of least action to the way modern supercomputers do simulations of reality to the way a scientist makes a discovery or an artist traps beauty on a canvass through the nuances of truth. The limit points of unconscious and super-conscious would correspond to the limiting cases of no information processing and infinite information processing respectively. Subconscious will be interpreted as partial information processing.

Every entity in the universe has to take a decision at every moment of time for its existence although the word existence may mean different things to different entities. The chance for continuation of existence is enhanced if the best decision on the basis of available information is taken. This is a process of optimization and the more conscious an entity is more is its ability to optimize.

Limitations of Understanding
Is there any fundamental principle (or, theorem) that puts limit on the understanding of both mind and matter by reducing them to information and then applying methodology of physical sciences to understand life and consciousness as emergent phenomena? Since this approach heavily relies on mathematics the limitations of deductive logic as pointed out by Gödel
in his famous incompleteness theorem may put the first limit. The second constraint may come from QT if it turns out (after having rigorous information theoretic formulation of both matter and mind) that the information related to mind is complementary to the information found in matter. I personally feel that this is quite unlikely because I believe that information at the fundamental level cannot be dualistic.

Conclusions

Unlike the Cartesian duality between mind and body, understanding consciousness requires first to understand matter and mind in a unified way. This can be achieved by giving information the most primary status in the universe. Then a generalized theory of quantum information dynamics has to be formulated (see the Table ??). The line of attack here involves three steps [Samal (2001)]:

1. understanding emergent phenomena and complexity in inanimate systems,
2. understanding life as emergent phenomena, and
3. understanding consciousness as emergent phenomena.

The attributes of consciousness can be understood only by a prudent application of both reductionism and holism. But the emergence of consciousness will be understood as an emergent phenomenon in the sense of structural organizations in the manifold of information to yield feasible structures through which we attribute meaning and understanding to the world.

7.1.7 Quantum Consciousness

For conscious states and brain states to mirror one another in any species, thereby establishing what von Neumann calls a psycho–physical parallelism, these intrinsically different states must evolve together and interact with one other during their time of evolution. Standard physics makes no provision for an interaction of this kind, but a quantum–mechanical opening for an objective/subjective interaction is shown to exist in [Mould (1995); Mould (1998); Mould (1999)]. In this subsection, following this approach, we present a model of quantum consciousness.

Our theory of subjective evolution calls for the existence of a Central Mechanism (CM) within an evolving organism, which contains presently unknown components of the nervous system. The function of a CM is to reduce quantum–mechanical superpositions within the nervous system,
and to simultaneously give rise to a conscious experience of the eigenvalues of the reduction. This accords with von Neumann’s requirement that a quantum–mechanical state reduction is accompanied by an observer’s conscious experience of the measured variables. At the present time, no one knows what there is about a conscious organism that gives rise to either consciousness or state reduction. We simply combined these two mysteries inside the CM, thereby placing our ignorance in a black–box so we can ask another question, namely: how do physical and mental states evolve interactively to insure the psycho–physical parallelism?

The model in [Mould (1995); Mould (1998); Mould (1999)] requires that a conscious organism spontaneously creates a profusion of macroscopic quantum–mechanical superpositions consisting of different neurological configurations. A mechanism for this generation is proposed by H. Stapp in [Stapp (1993)]. The result is a superposition of different neurological states, each of which may be accompanied by a different subjective experience. A reduction to a single eigenstate is not assumed to be triggered microscopically along the lines of [Ghirardi et al. (1986)]; but rather, it is assumed to occur in response to a macroscopic event. It occurs the moment an emerging subjective state becomes actively conscious in one of the macroscopic neurological components of a Stapp superposition. The consciousness that is associated with such a reduction is assumed to fade the moment reduction is complete, and the resulting subjective pulse is supposedly followed by similar pulses in rapid succession. This can make the subject aware of an apparent continuum of consciousness.

Presumably, any reduction of this kind is accompanied by a reduction of all other parts of the organism as well as all those parts of the external world that are correlated with it. This means that a second observer, coming on the heels of the first, will make an observation in agreement with the first. More formally, a measurement interaction establishes correlations between the eigenstates \( |a_i\rangle \) of some apparatus (with discrete variables \( a_i \)), eigenstates of a first observer \( |\Phi_i\rangle \), and eigenstates of a second observer \( |\Theta_i\rangle \), such that the total state prior to reduction is given by [Mould (1995); Mould (1998); Mould (1999)]

\[
|\Psi\rangle = \sum_i C_i |a_i\rangle |\Phi_i\rangle |\Theta_i\rangle.
\]

The coefficient \( C_i \) is the probability amplitude that the apparatus is in state \( |a_i\rangle \). Let the first observer become consciously aware of the apparatus variable \( a_k \). The resulting reduction is a projection in Hilbert space that is found by applying the projection operator of that observer \( |\Phi_k\rangle \langle \Phi_k| \) to the
Let the second observer then become consciously aware of the apparatus variable $a_m$. The subsequent reduction is found by applying the projection operator of that observer $|\Phi_m\rangle\langle\Phi_m|$ to the first reduction.

$$|\Theta_m\rangle\langle\Theta_m|C_k|a_k\rangle|\Phi_k\rangle|\Theta_k\rangle = \delta_{km}C_k|a_k\rangle|\Phi_k\rangle|\Theta_m\rangle$$

(2nd reduction)

Only if $m = k$ is the probability non–zero that the second observer will make a measurement. The second observer therefore confirms the results of the first observer that the apparatus has been left in the eigenstate $|a_k\rangle$.

Again, many of the particulars of a reduction (such as its nonlinearity) are ignored in this subsection so we can concentrate on the influence of subjective states on physiological states. To this end we require that ‘when the emerging subjective states of a neurological superposition are different from one another, they will generally exert an influence on their relative probability amplitudes that is a function of that difference’ [Mould (1995); Mould (1998); Mould (1999)]. In particular, we imagine that when a ‘painful’ subjective state emerges in superposition with a ‘pleasurable’ subjective state, the probability amplitude of the painful state will be decreased relative to the probability amplitude of the pleasurable state.

No currently known observation contradicts this conjecture, for no previously reported experiment deals specifically with the creation of different observers experiencing different degrees of pain, arising on different components of a quantum mechanical superposition.

Let $N$ in Figure 7.4 represent the nervous system of the first primitive organism that makes a successful use of the subjective experience of ‘pain’. In [Mould (1995); Mould (1998); Mould (1999)] this creature was imagined to be a fish. It is supposed that the fish makes contact with an electric probe, at which time its nervous system splits into a superposition (via the Stapp mechanism) consisting of a withdrawal behavior $W$ that is accompanied by [no pain], and a continued contact behavior $C$ that is accompanied by [pain]. The probability of survival of each component in this highly artificial model is initially assumed to be 0.5. However, because of the hypothetical influence of subjective pain on probability amplitudes, only the withdrawal state is assumed to survive the reduction in this idealized example. State reduction in Figure 7.4 is represented by the horizontal arrow. If $W$ is furthermore a good survival strategy from the point of view
of evolution, then the association $W[\text{no pain}]$ and $C[\text{pain}]$ will serve the species well, whereas a wrong association $W[\text{pain}]$ and $C[\text{no pain}]$ will lead to its demise.

It does not matter to the above argument if the variables are ‘pleasure/pain’ or some other range of subjective experiences. If a subjective experience like ‘$A$’ increases the probability amplitude of an escape behavior, and if a subjective experience like ‘$B$’ diminishes the probability amplitude of that behavior, and if the escape is one that moves the creature away from something that is dangerous to its health, then a distant descendent will experience ‘$A$’ associated with life supporting escapes, and ‘$B$’ associated with life threatening failures-to-escape. It is apparent that the quality of the experience does not matter. We require only that the subjective experience in question has a predictable plus or minus effect on the probability amplitudes within a superposition, and the survival mechanisms of evolution will do the rest. They will insure that the eventual subjective life of a surviving species mirrors its experiences in a definite and predictable way thereby establishing a reliable psycho-physical parallelism.

We assume that ordinary perception do not have this effect. They do not give rise to the hypothetical feedback. In Figure 7.5 we imagine the existence of an externally imposed two component superposition consisting of environments $e_1$ and $e_2$, which is produced by using, say, a $\beta$ source. The two environments are assumed to have equal probability, and are allowed to interact with the subject’s nervous system given by $N_0$. Before a reduction can occur, two conscious states emerge from the interaction represented by the superposition of $(eN_1)[x_1]$ and $(eN_2)[x_2]$, where the conscious part shown in brackets is the observed eigenvalue $x$ associated with components 1
and 2. Since we require that an observer of the ‘perceived’ variable $x$ cannot affect the probability of $x$, the pure state reduces to a mixture having the same probability as the initial superposition (horizontal arrows in Figure 7.5). State $e_1$ represents the relevant laboratory apparatus together with the wider environment with which it is entangled. The phase angles $\phi$ and $\phi'$ are definite, but they are not localized to manageable parts of the apparatus [Mould (1995); Mould (1998); Mould (1999)]. We call them ‘arbitrary’ to indicate that their values are not practically calculable, and to emphasize the lack of coherence between these ‘macroscopic’ components.

On the other hand, if ‘pain’ were the variable in Figure 7.5 rather than the externally perceived variable $x$, it is suggested by the hypothesis of [Mould (1995); Mould (1998); Mould (1999)] that the resulting mixture might no longer be a 50–50 split. This possibility is represented in Figure 7.6, where the final mixture probabilities are left unspecified because they must be discovered by observation.

7.1.7.1 The Experiment

Two scalers L and R recording local background radiation are placed side-by-side in Figure 7.7. Their outputs are fed to a selector box that chooses channels L or R, depending on which is the first to record a single count after the selector has been turned on. A 20 V signal is then emitted from the output of the chosen channel. The output on the R-channel is unused, but the L-output closes a relay that puts 80 volts across two metal bars. Two seconds after the selection, an L or R-light goes on indicating which channel was selected. A finger placed across the metal bars will receive a painful 80V shock when the L-channel is selected [Mould (1995); Mould (1998); Mould (1999)].

This apparatus allows us to carry out the experiments diagramed in Figures 7.5 and 7.6. If the selector is initiated in the absence of an observer, we say that the system will become a macroscopic superposition given by $(\cos \phi e_1 + e_2)$, where $e_1$ is the entire apparatus following an L-channel activation, and $e_2$ is the entire apparatus following an R-channel activation.
The incoherence of the two components (represented by the arbitrary angle $\phi$) is generally understood to mean that the system is indistinguishable from a classical mixture, since interference between these macroscopic components is not possible. However, for reasons given in previous papers, we claim that the final state is really an incoherent quantum–mechanical superposition rather than a classical mixture. The lack of interference between the components has no bearing on our result because the hypothetical effect described here relates to, and directly affects, probability amplitudes only. The effect we are looking for should be observable with or without coherence between L and R.

Fig. 7.7 The experimental set up (see text for explanation).

If an observer is present and exposed only to the L–light or the R–light, then a reduction will occur like the one in Figure 7.5 where eigenvalues $x_1$ and $x_2$ represent a conscious experience of one or the other of those lights. If the observer is exposed only to a conscious experience of “pain or no pain” through his finger across the metal bars, then a reduction like the one in Figure 7.6 will occur. This experiment may not appear to

---

$^{34}$The uncertainty associated with a classical mixture state represents an outsider’s ignorance, whereas a pure quantum mechanical state superposition represents an uncertainty that is intrinsic to the system. Following von Neumann, we assume that the initial intrinsic uncertainty (concerning which of the scalers fires first) will remain an intrinsic uncertainty until it is reduced by ‘observation’. Hence, the apparatus will remain a macroscopic pure state quantum mechanical superposition until an observation occurs.
be quantum–mechanical, but it is quantum–mechanical by virtue of the particular hypothesis that is being tested in Figure 7.6.

The equipment in Figure 7.7 was used for a total of 2500 trials, each consisting of two parts. The experimenter’s finger was first placed across the metal bars, the selector was turned on, and a ‘shock’ or ‘no shock’ was recorded before the lights were observed. In the second part of each trial the finger was replaced by an equivalent resistance, the selector was again initiated, and the appearance of the L or R channel light was recorded [Mould (1995); Mould (1998); Mould (1999)].

\[ N = 2500, \]
\[ N_S = 1244, \]
\[ N_L = 1261. \]

There are three possible outcomes of a single trial. Either the difference \( N_L - N_S \) increases, or it decreases, or it remains the same. The three possibilities are represented by the variables \( u \) (increase) occurring with a probability \( p \), and \( d \) (decrease) with a probability \( q \), and \( e \) (remain the same) with a probability \( r \). It was found in the experiment that \( u = 632 \) and \( d = 615 \) after 2500 trials.

If we approximate \( p_0 = N_L/N \) to be the probability that the left channel fires in the second part of each trial (absent the finger), and \( q_0 = 1 - p_0 \) to be the probability that the right channel fires in the second part of each trial, then

\[ p_0 = 1261/2500 = 0.5044 \quad q_0 = 0.4956 \]

Assuming as a null hypothesis that there is no statistical difference between the displacement of a finger across the metal bars and an equivalent resistor, we have \( p = p_0 q_0 \), \( q = q_0 p_0 \), and \( r = p_0^2 + q_0^2 \), giving

\[ p = 0.2500 \quad q = 0.2500 \quad r = 0.5000 \]

The variances of \( (u + d) \) and \( (u - d) \) are

\[ \sigma^2(u + d) = \langle (u + d)^2 \rangle - (u + d)^2 = \sigma^2(u) + \sigma^2(d) + X \]
\[ \sigma^2(u - d) = \langle (u - d)^2 \rangle - (u - d)^2 = \sigma^2(u) + \sigma^2(d) - X \]
therefore

\[ \sigma^2(u - d) = 2\sigma^2(u) + 2\sigma^2(d) - \sigma^2(u + d) \]
\[ = 2p(q + r)N + 2q(p + r)N - r(p + q)N \]

or

\[ \sigma(u - d) = \left[4pq + r(p + q)\right]N^{1/2} = \left[N/2\right]^{1/2} = 35.4 \]

Our alternative hypothesis is that \( u - d \) is significantly different from 0. But from the data, \( u - d = N_L - N_S = 17 \) after 2500 trials, and this is well within the above the standard deviation around 0. The separate variables \( u \) and \( d \) are also within the standard deviation

\[ \sigma(u) = \sigma(d) = \left[p(q + r)N\right]^{1/2} = 21.7 \]

of their expected value of 625.

One can always argue that the statistics are inadequate to reveal a significant difference between \( u \) and \( d \). However, they are sufficient to convince us that the presence of pain on one component of this externally imposed superposition has no significant effect on the outcome. We therefore conclude that the reduction in Figure 7.6 is not affected by the subjective content of the square brackets in that figure.

7.1.7.2 Bio–Active Peptides

Neurological communication depends on the diffusion of chemical neurotransmitters across the synaptic junction between neurons. There is another communication system within the body that makes use of chemicals that are produced at one site and received at another; but in this case, the distances between a production and receiver sites are macroscopic. About 95% of these chemical communicators are peptides, which are mini–proteins consisting of up to 100 amino acids having a maximum atomic mass of 10,000 u. Their classical dimensions are \( \Delta x = 10 \) nm at most, which we assume approximates their size close to the production site [Pert (1997)]. Therefore, Heisenberg tells us that the minimum quantum–mechanical uncertainty in the velocity of one of these free peptides is \( \Delta v = 0.63 \) mm/s. Peptides are carried through intercellular space by blood and cerebrospinal fluid. They do not move very far in a tenth of a second, but in that time the Heisenberg uncertainty in position of a peptide will be at least \( \Delta s = \Delta v \Delta t = 63 \) mm. This is an enormous uncertainty of position relative to one of the peptide receptor sites which has a size similar to that of the peptide, and which is often separated from its neighbors by comparable distances. Therefore, quantum–mechanical
uncertainty is an important factor in determining the probability that a given peptide is captured by a given receptor \[\text{Mould (1995); Mould (1998); Mould (1999)}\].

Stapp’s mechanism for introducing quantum–mechanical superpositions into the brain relies on the uncertainty in the position of calcium ions in neuron synapses. We suggest that peptides represent another possible source of super-positions that may be just as widespread. And because peptides play an important role in the chemistry of the body, they too may have a significant quantum–mechanical influence on behavior.

As with the Stapp mechanism, one might object that the uncertainty associated with the peptide’s classical diffusion during its migration will overwhelm the quantum–mechanical uncertainty, or that a large number of migrating molecules will obscure all quantum–mechanical effects. However, the classical uncertainty associated with many–particle ensembles has only to do with our ignorance of initial conditions. In reality, the only uncertainties a receptor will see are those associated with an incoherent quantum–mechanical superposition of pure peptide states. This superposition will have as many components as there are peptide molecules involved. And since our hypothetical influence acts through the amplitude of these components, the presence of a large number of independent particles will only increase the hypothetical influence.

7.1.7.3 Drugs
There are many drugs that can be introduced into the body that will compete with endogenous peptides to occupy the body’s receptor sites, and some of these drug molecules are small enough to have a very large quantum mechanical uncertainty of position. For this reason, peptide/drug superpositions are more promising for the purpose of experimental manipulation than calcium ion super-positions.

For example, endorphins are peptides that unite with special receptors to eliminate pain and/or produce euphoria. They and their receptors can be found everywhere in the body, but they are most intensely located in the limbic system of the brain. There is a drug called naloxone that is a strong competitor with the endorphins to occupy the same receptors, and it has the property that it reverses the analgesic/pleasurable effects of the endorphins \[\text{Pert (1997); Snyder (1986); Levinthal (1988)}\]. If endorphin molecules and externally administered naloxone molecules are in quantum–mechanical superposition with one another
as their sizes and likely time together suggests, and if they both compete
with one another for successful attachment to the same receptor site, then
the ratio of endorphin attachments to naloxone attachments would (accord-
ing to our hypothesis) be a function of the competing subjective states.

7.1.7.4 Evolutionary Advantage

It was pointed out in [Mould (1995); Mould (1998); Mould (1999)] that our
evolutionary mechanism of objective–subjective interaction (represented by
Figure 7.4) does not insure that a creature evolving under its influence will
evolve more quickly or be more successful than a creature evolving strictly
as an automaton. That will be true as well of the modified model in sects.
3-5. However, it is not unreasonable to suppose that both conscious evo-
lution and autonomic evolution might work separately and in tandem with
one another. The kinds of neurological changes that are necessary for auto-
nomic evolution might very well be independent of the kinds of neurological
changes that are necessary for quantum/consciousness evolution. If that is
so, and if these two processes work in tandem, then the evolution of the
organism will be faster than either the autonomic route by itself, or the
conscious route by itself. One would then be able to say that the introduc-
tion of consciousness as proposed here will always work to the advantage
of the organism.

7.1.8 Quantum–Like Psychodynamics

In this section, which is written in the fashion of the quantum brain,
we present the top level of natural biodynamics, using geometrical gen-
eralization of the Feynman path integral. To formulate the basics of
force–field psychodynamics, we use the action–amplitude picture of the
BODY ⇔ MIND adjunction [Ivancevic and Ivancevic (2007c); Ivance-
vic and Ivancevic (2007d)]:

↓ Deterministic (causal) world of Human BODY ↓

Action : \( S[q^n] = \int_{t_{in}}^{t_{out}} \left( E_k - E_p + Wrk + Src^\pm \right) dt \)

Amplitude : \( \langle \text{out}|\text{in} \rangle = \oint \mathcal{D}[w_nq^n] e^{iS[q^n]} \)

↑ Probabilistic (fuzzy) world of Human MIND ↑
In the action integral, $E_k, E_p, W_{rk}$ and $S_{rc}$ denote the kinetic end potential energies, work done by dissipative/driving forces and other energy sources/sinks, respectively. In this way, the whole body dynamics is incorporated in the mind dynamics. This adaptive path integral represents an infinite–dimensional neural network, suggesting an infinite capacity of human brain/mind.

For a long time the cortical systems for language and actions were believed to be independent modules. However, according to the recent research of [Pulvermüller (2005)], as these systems are reciprocally connected with each other, information about language and actions might interact in distributed neuronal assemblies. A critical case is that of action words that are semantically related to different parts of the body (e.g. ‘pick’, ‘kick’, ‘lick’,...). The author suggests that the comprehension of these words might specifically, rapidly and automatically activate the motor system in a somatotopic manner, and that their comprehension rely on activity in the action system.

7.1.8.1 Life Space Foam

Applications of nonlinear dynamical systems in psychology have been encouraging, if not universally effective [Metzger (1997)]. Its historical antecedents can be traced back to Piaget’s [Piaget et al. (1992)] and Vygotsky’s [Vygotsky (1982)] interpretations of the dynamic relations between action and thought, Lewin’s theory of social dynamics and cognitive–affective development [Lewin (1997)], and [Bernstein (1947)] theory of self–adjusting, goal–driven motor action.

Now, both the original Lewinian force–field theory in psychology (see [Lewin (1951); Gold (1999)]) and modern decision–field dynamics (see [Busemeyer and Townsend (1993); Roe et al. (2001); Busemeyer and Diederich (2002)]) are based on the classical Lewinian concept of an individual’s life space.\[^{35}\] As a topological construct, Lewinian life space represents a person’s psychological environment that contains regions separated by dynamical permeable boundaries. As a field construct, on the other hand, the life space is not empty: each of its regions is characterized by valence (ranging from positive or negative and resulting from an interaction between the person’s needs and the dynamics of their environment). Need is an energy

\[^{35}\]The work presented in this subsection has been developed in collaboration with Dr. Eugene Aidman, Senior Research Scientist, Human Systems Integration, Land Operations Division, Defence Science & Technology Organisation, Australia.
It creates tension in the person, which, in combination with other tensions, initiates and sustains behavior. Needs vary from the most primitive urges to the most idiosyncratic intentions and can be both internally generated (e.g., thirst or hunger) and stimulus–induced (e.g., an urge to buy something in response to a TV advertisement). Valences are, in essence, personal values dynamically derived from the person’s needs and attached to various regions in their life space. As a field, the life space generates forces pulling the person towards positively–valenced regions and pushing them away from regions with negative valence. Lewin’s term for these forces is vectors. Combinations of multiple vectors in the life space cause the person to move from one region towards another. This movement is termed locomotion and it may range from overt behavior to cognitive shifts (e.g., between alternatives in a decision–making process). Locomotion normally results in crossing the boundaries between regions. When their permeability is degraded, these boundaries become barriers that restrain locomotion. Life space model, thus, offers a meta–theoretical language to describe a wide range of behaviors, from goal–directed action to intrapersonal conflicts and multi–alternative decision–making.

In order to formalize the Lewinian life–space concept, a set of action principles need to be associated to Lewinian force–fields, (loco)motion paths (representing mental abstractions of biomechanical paths [Ivancevic and Ivancevic (2005)]) and life space geometry. As an extension of the Lewinian concept, in this section we recall [Ivancevic and Aidman (2007)] a new concept of life–space foam (LSF, see Figure 7.8). According to this new concept, Lewin’s life space can be represented as a geometrical functor with globally smooth macro–dynamics, which is at the same time underpinned by wildly fluctuating, non–smooth, local micro–dynamics, describable by Feynman’s: (i) sum–over–histories $\int \Sigma$ paths, (ii) sum–over–fields $\int \Sigma$ fields, and (iii) sum–over–geometries $\int \Sigma$ geom.

LSF is thus a two–level geometrodynamical functor, representing these two distinct types of dynamics within the Lewinian life space. At its macroscopic spatio–temporal level, LSF appears as a ‘nice & smooth’ geometrical functor with globally predictable dynamics – formally, a smooth $n$–dimensional manifold $M$ with local Riemannian metrics $g_{ij}(x)$, smooth force–fields and smooth (loco)motion paths, as conceptualized in the Lewinian theory. To model the global and smooth macro–level LSF–paths, fields and geometry, we use the general physics–like principle of the least action.
Now, the apparent smoothness of the macro-level LSF is achieved by the existence of another level underneath it. This micro-level LSF is actually a collection of wildly fluctuating force-fields, (loc)motion paths, curved regional geometries and topologies with holes. The micro-level LSF is proposed as an extension of the Lewinian concept: it is characterized by uncertainties and fluctuations, enabled by microscopic time-level, microscopic transition paths, microscopic force-fields, local geometries and varying topologies with holes. To model these fluctuating microscopic LSF-structures, we use three instances of adaptive path integral, defining a multi-phase and multi-path (also multi-field and multi-geometry) transition process from intention to the goal-driven action.

We use the new LSF concept to develop modelling framework for motivational dynamics (MD) and induced cognitive dynamics (CD).

According to Heckhausen (see Heckhausen (1977)), motivation can be thought of as a process of energizing and directing the action. The process of energizing can be represented by Lewin’s force-field analysis and Vygotsky’s motive formation (see Vygotsky (1982); Aidman and Leontiev (1991)), while the process of directing can be represented by hierarchical action control (see Bernstein (1947); Bernstein (1935); Kuhl (1985)).

Motivation processes both precede and coincide with every goal-directed action. Usually these motivation processes include the sequence of the following four feedforward phases Vygotsky (1982); Aidman and Leontiev (1991): (*)

1. **Intention Formation** $F$, including: decision making, commitment building, etc.
2. **Action Initiation** $I$, including: handling conflict of motives, resistance to alternatives, etc.
Fig. 7.9 Transition–propagator corresponding to each of the motivational phases \( \{F, I, M, T\} \), consisting of an ensemble of feedforward paths propagating through the ‘wood of obstacles’. The paths affected by driving and restraining force–fields, as well as by the local LSF–geometry. Transition goes from Intention, occurring at a sample time instant \( t_0 \), to Action, occurring at some later time \( t_1 \). Each propagator is controlled by its own Monitor feedback. All together they form the transition functor \( TA \).

(3) Maintaining the Action \( M \), including: resistance to fatigue, distractions, etc.

(4) Termination \( T \), including parking and avoiding addiction, i.e., staying in control.

With each of the phases \( \{F, I, M, T\} \) in (*), we can associate a transition propagator — an ensemble of (possibly crossing) feedforward paths propagating through the ‘wood of obstacles’ (including topological holes in the LSF, see Figure 7.9), so that the complete transition functor \( TA \) is a product of propagators (as well as sum over paths). All the phases–propagators are controlled by a unique Monitor feedback process.

In this subsection we propose an adaptive path integral formulation for the motivational–transition functor \( TA \). In essence, we sum/integrate over different paths and make a product (composition) of different phases–propagators. Recall that this is the most general description of the general Markov stochastic process.

We will also attempt to demonstrate the utility of the same LSF–formalisms in representing cognitive functions, such as memory, learning and decision making. For example, in the classical Stimulus encoding →
Search → Decision → Response sequence \cite{Sternberg(1969), Ashcraft(1994)}, the environmental input–triggered sensory memory and working memory (WM) can be interpreted as operating at the micro–level force–field under the executive control of the Monitor feedback, whereas search can be formalized as a control mechanism guiding retrieval from the long–term memory (LTM, itself shaped by learning) and filtering material relevant to decision making into the WM. The essential measure of these mental processes, the processing speed (essentially determined by Sternberg’s reaction–time) can be represented by our (loco)motion speed $\dot{x}$.

**Six Faces of the Life Space Foam**

The LSF has three forms of appearance: paths + field + geometries, acting on both macro–level and micro–level, which is six modes in total. In this section, we develop three least action principles for the macro–LSF–level and three adaptive path integrals for the micro–LSF–level. While developing our psycho–physical formalism, we will address the behavioral issues of motivational fatigue, learning, memory and decision making.

**General Formalism**

At both macro– and micro–levels, the total LSF represents a union of transition paths, force–fields and geometries, formally written as

$$LSF_{\text{total}} := LSF_{\text{paths}} \cup LSF_{\text{fields}} \cup LSF_{\text{geom}}$$

$$\equiv \oint \Sigma \text{paths} + \oint \Sigma \text{fields} + \oint \Sigma \text{geom}.$$  \hspace{1cm} (7.35)

Corresponding to each of the three LSF–subspaces in (7.35) we formulate:

1. The least action principle, to model deterministic and predictive, macro–level MD & CD, giving a unique, global, causal and smooth path–field–geometry on the macroscopic spatio–temporal level; and

2. Associated adaptive path integral to model uncertain, fluctuating and probabilistic, micro–level MD & CD, as an ensemble of local paths–fields–geometries on the microscopic spatio–temporal level, to which the global macro–level MD & CD represents both time and ensemble average (which are equal according to the ergodic hypothesis).

In the proposed formalism, transition paths $x^i(t)$ are affected by the force–fields $\varphi^k(t)$, which are themselves affected by geometry with metric $g_{ij}$. 

---

*Quantum Mind* 743
Global Macro–Level of $LSF_{\text{total}}$. In general, at the macroscopic LSF–level we first formulate the total action $S[\Phi]$, the central quantity in our formalism that has psycho–physical dimensions of $\text{Energy} \times \text{Time} = \text{Effort}$, with immediate cognitive and motivational applications: the greater the action – the higher the speed of cognitive processes and the lower the macroscopic fatigue (which includes all sources of physical, cognitive and emotional fatigue that influence motivational dynamics). The action $S[\Phi]$ depends on macroscopic paths, fields and geometries, commonly denoted by an abstract field symbol $\Phi^i$. The action $S[\Phi]$ is formally defined as a temporal integral from the initial time instant $t_{\text{ini}}$ to the final time instant $t_{\text{fin}}$,

$$S[\Phi] = \int_{t_{\text{ini}}}^{t_{\text{fin}}} \mathcal{L}[\Phi] \, dt,$$

with Lagrangian density given by

$$\mathcal{L}[\Phi] = \int d^n x \, \mathcal{L}(\Phi^i, \partial_x \Phi^i),$$

where the integral is taken over all $n$ coordinates $x^i = x^i(t)$ of the LSF, and $\partial_x \Phi^i$ are time and space partial derivatives of the $\Phi^i$–variables over coordinates.

Second, we formulate the least action principle as a minimal variation $\delta$ of the action $S[\Phi]$

$$\delta S[\Phi] = 0,$$

which, using techniques from the calculus of variations gives, in the form of the so–called Euler–Lagrangian equations, a shortest (loco)motion path, an extreme force–field, and a life–space geometry of minimal curvature (and without holes). In this way, we effectively derive a unique globally smooth transition functor

$$T \, A : INTENTION_{t_{\text{ini}}} \Rightarrow ACTION_{t_{\text{fin}}},$$

performed at a macroscopic (global) time–level from some initial time $t_{\text{ini}}$ to the final time $t_{\text{fin}}$.

In this way, we get macro–objects in the global LSF: a single path described Newtonian–like equation of motion, a single force–field described by Maxwellian–like field equations, and a single obstacle–free Riemannian geometry (with global topology without holes).
For example, recall that in the period 1945–1949, John Wheeler and Richard Feynman developed their action–at–a–distance electrodynamics \cite{WheelerFeynman}, in complete experimental agreement with the classical Maxwell’s electromagnetic theory, but at the same time avoiding the complications of divergent self–interaction of the Maxwell’s theory as well as eliminating its infinite number of field degrees–of–freedom. In Wheeler–Feynman view, “Matter consists of electrically charged particles,” so they found a form for the action directly involving the motions of the charges only, which upon variation would give the Newtonian–like equations of motion of these charges. Here is the expression for this action in the flat space–time, which is in the core of quantum electrodynamics:

\[
S[x; t_i, t_j] = \frac{1}{2} m_i \int (\dot{x}^i_\mu)^2 \, dt_i + \frac{1}{2} e_i e_j \int \delta(I^2_{ij}) \dot{x}^i_\mu(t_i) \dot{x}^j_\mu(t_j) \, dt_i dt_j
\]

with

\[
I^2_{ij} = \left[ x^i_\mu(t_i) - x^j_\mu(t_j) \right] \left[ x^i_\mu(t_i) - x^j_\mu(t_j) \right],
\]

where \( x^i_\mu = x^i_\mu(t_i) \) is the four–vector position of the \( i \)th particle as a function of the proper time \( t_i \), while \( \dot{x}^i_\mu(t_i) = dx^i_\mu/dt_i \) is the velocity four–vector. The first term in the action (7.39) is the ordinary mechanical action in Euclidean space, while the second term defines the electrical interaction of the charges, representing the Maxwell–like field (it is summed over each pair of charges; the factor \( \frac{1}{2} \) is to count each pair once, while the term \( i = j \) is omitted to avoid self–action; the interaction is a double integral over a delta function of the square of space–time interval \( I^2 \) between two points on the paths; thus, interaction occurs only when this interval vanishes, that is, along light cones \cite{WheelerFeynman}).

Now, from the point of view of Lewinian geometrical force–fields and (loco)motion paths, we can give the following life–space interpretation to the Wheeler–Feynman action (7.39). The mechanical–like locomotion term occurring at the single time \( t_i \), needs a covariant generalization from the flat 4D Euclidean space to the \( n \)D smooth Riemannian manifold, so it becomes (see e.g., \cite{IvancevicIvancevic})

\[
S[x] = \frac{1}{2} \int_{t_{ini}}^{t_{fin}} g_{ij} \dot{x}^i \dot{x}^j \, dt,
\]

where \( g_{ij} \) is the Riemannian metric tensor that generates the total ‘kinetic energy’ of (loco)motions in the life space.

The second term in (7.39) gives the sophisticated definition of Lewinian force–fields that drive the psychological (loco)motions, if we interpret elec-
trical charges \(e_i\) occurring at different times \(t_i\) as motivational charges – needs.

**Local Micro–Level of \(LSF_{total}\).** After having properly defined macro–level MD & CD, with a unique transition map \(F\) (including a unique motion path, driving field and smooth geometry), we move down to the microscopic LSF–level of rapidly fluctuating MD & CD, where we cannot define a unique and smooth path–field–geometry. The most we can do at this level of fluctuating uncertainty, is to formulate an adaptive path integral and calculate overall probability amplitudes for ensembles of local transitions from one LSF–point to the neighboring one. This probabilistic transition micro–dynamics functor is defined by a multi–path (field and geometry, respectively) and multi–phase transition amplitude \(\langle \text{Action}|\text{Intention} \rangle\) of corresponding to the globally–smooth transition map \((7.38)\). This absolute square of this probability amplitude gives the transition probability of occurring the final state of \(\text{Action}\) given the initial state of \(\text{Intention}\),

\[
P(\text{Action}|\text{Intention}) = |\langle \text{Action}|\text{Intention} \rangle|^2.
\]

The total transition amplitude from the state of \(\text{Intention}\) to the state of \(\text{Action}\) is defined on \(LSF_{total}\)

\[
TA \equiv \langle \text{Action}|\text{Intention} \rangle_{total} : \text{INTENTION}_{t_0} \Rightarrow \text{ACTION}_{t_1}, \quad (7.40)
\]
given by adaptive generalization of the Feynman’s path integral [Ivancevic and Ivancevic (2007b)]. The transition map \((7.40)\) calculates the overall probability amplitude along a multitude of wildly fluctuating paths, fields and geometries, performing the microscopic transition from the micro–state \(\text{INTENTION}_{t_0}\) occurring at initial micro–time instant \(t_0\) to the micro–state \(\text{ACTION}_{t_1}\) at some later micro–time instant \(t_1\), such that all micro–time instants fit inside the global transition interval \(t_0, t_1, ..., t_s \in [t_{ini}, t_{fin}]\).

It is symbolically written as

\[
\langle \text{Action}|\text{Intention} \rangle_{total} := \sum_{\Phi} \mathcal{D}[w\Phi] e^{iS[\Phi]}, \quad (7.41)
\]

where the Lebesgue integration is performed over all continuous \(\Phi_{con} = \text{paths} + \text{field} + \text{geometries}\), while summation is performed over all discrete processes and regional topologies \(\Phi_{dis}^j\). The symbolic differential \(\mathcal{D}[w\Phi]\) in the general path integral \((7.70)\), represents an adaptive path measure, defined as a weighted product

\[
\mathcal{D}[w\Phi] = \lim_{\lambda \to -\infty} \prod_{s=1}^{N} w_s d\Phi_s^i, \quad (i = 1, ..., n = con + dis), \quad (7.42)
\]
which is in practice satisfied with a large $N$ corresponding to infinitesimal
temporal division of the four motivational phases (*). Technically, the path
integral (7.70) calculates the amplitude for the transition functor $\mathcal{T}A :$
Intention $\Rightarrow$ Action.

In the exponent of the path integral (7.70) we have the action $S[\Phi]$ and
the imaginary unit $i = \sqrt{-1}$ ($i$ can be converted into the real number $-1$
using the so-called Wick rotation, see next subsection).

In this way, we get a range of micro–objects in the local LSF at the
short time–level: ensembles of rapidly fluctuating, noisy and crossing paths,
force–fields, local geometries with obstacles and topologies with holes. However,
by averaging process, both in time and along ensembles of paths, fields
and geometries, we recover the corresponding global MD & CD variables.

**Infinite–Dimensional Neural Network.** The adaptive path integral
(7.70) incorporates the local learning process according to the standard
formula: New Value $= \text{Old Value} + \text{Innovation}$. The general weights $w_s = w_s(t)$ in (7.42) are updated by the MONITOR feedback during the
transition process, according to one of the two standard neural learning
schemes, in which the micro–time level is traversed in discrete steps, i.e., if
$t = t_0, t_1, ..., t_s$ then $t + 1 = t_1, t_2, ..., t_{s+1}$:

1. A self–organized, unsupervised (e.g., Hebbian–like [Hebb (1949)]) learning rule:

$$w_s(t + 1) = w_s(t) + \frac{\sigma}{\eta}(w_s^d(t) - w_s^a(t)), \quad (7.43)$$

where $\sigma = \sigma(t)$, $\eta = \eta(t)$ denote signal and noise, respectively, while
superscripts $d$ and $a$ denote desired and achieved micro–states, respectively; or

2. A certain form of a supervised gradient descent learning:

$$w_s(t + 1) = w_s(t) - \eta \nabla J(t), \quad (7.44)$$

where $\eta$ is a small constant, called the step size, or the learning rate
and $\nabla J(n)$ denotes the gradient of the ‘performance hyper–surface’ at
the $t$–th iteration.

Both Hebbian and supervised learning are used for the local decision making
process (see below) occurring at the intention formation phase $\mathcal{F}$.

In this way, local micro–level of $\text{LSF}_{\text{total}}$ represents an infinite–dimensional neural network. In the cognitive psychology framework, our
adaptive path integral (7.70) can be interpreted as semantic integration (see Bransford and Franks (1971) [Ashcraft (1994)]).

**Motion and Decision Making in LSF**

On the macro–level in the subspace LSF, we have the (loco)motion action principle

$$\delta S[x] = 0,$$

with the Newtonian–like action $S[x]$ given by

$$S[x] = \int_{t_{ini}}^{t_{fin}} dt \left[ \frac{1}{2} g_{ij} \dot{x}^i \dot{x}^j + \varphi^i(x^i) \right], \tag{7.45}$$

where overdot denotes time derivative, so that $\dot{x}^i$ represents processing speed, or (loco)motion velocity vector. The first bracket term in (7.45) represents the kinetic energy $T$,

$$T = \frac{1}{2} g_{ij} \dot{x}^i \dot{x}^j,$$

generated by the Riemannian metric tensor $g_{ij}$, while the second bracket term, $\varphi^i(x^i)$, denotes the family of potential force–fields, driving the (loco)mo-tions $x^i = x^i(t)$ (the strengths of the fields $\varphi^i(x^i)$ depend on their positions $x^i$ in LSF, see LSF fields below). The corresponding Euler–Lagrangian equation gives the Newtonian–like equation of motion

$$\frac{d}{dt} T_{x^i} - T_{x^i} = -\varphi_{x^i}, \tag{7.46}$$

(subscripts denote the partial derivatives), which can be put into the standard Lagrangian form

$$\frac{d}{dt} L_{x^i} = L_{x^i}, \quad \text{with} \quad L = T - \varphi^i(x^i).$$

In the next subsection we use the micro–level implications of the action $S[x]$ as given by (7.45), for dynamical descriptions of the local decision–making process.

On the micro–level in the subspace LSF, instead of a single path defined by the Newtonian–like equation of motion (7.46), we have an ensemble of fluctuating and crossing paths with weighted probabilities (of
the unit total sum). This ensemble of micro–paths is defined by the simplest instance of our adaptive path integral \((7.70)\), similar to the Feynman’s original sum over histories,

\[
\langle \text{Action}|\text{Intention}\rangle_{\text{paths}} = \oint D[wx] e^{iS[x]},
\]

(7.47)

where \(D[wx]\) is a functional measure on the space of all weighted paths, and the exponential depends on the action \(S[x]\) given by \((7.45)\). This procedure can be redefined in a mathematically cleaner way if we Wick–rotate the time variable \(t\) to imaginary values \(t \mapsto \tau = it\), thereby making all integrals real:

\[
\oint D[wx] e^{iS[x]} \Rightarrow \text{Wick} \oint D[wx] e^{-S[x]}.
\]

(7.48)

Discretization of \((7.48)\) gives the thermodynamic–like partition function

\[
Z = \sum_j e^{-w_j E_j/T},
\]

(7.49)

where \(E_j\) is the motion energy eigenvalue (reflecting each possible motivational energetic state), \(T\) is the temperature–like environmental control parameter, and the sum runs over all motion energy eigenstates (labelled by the index \(j\)). From \((7.49)\), we can further calculate all thermodynamic–like and statistical properties of MD & CD (see e.g., \([\text{Feynman (1972)}]\)), as for example, transition entropy \(S = k_B \ln Z\), etc.

From cognitive perspective, our adaptive path integral \((7.47)\) calculates all (alternative) pathways of information flow during the transition \(\text{Intention} \rightarrow \text{Action}\).

In the language of transition–propagators, the integral over histories \((7.47)\) can be decomposed into the product of propagators (i.e., Fredholm kernels or Green functions) corresponding to the cascade of the four motivational phases (*):

\[
\langle \text{Action}|\text{Intention}\rangle_{\text{paths}} = \oint dx^F dx^T dx^M dx^T K(F,T)K(I,M)K(M,T),
\]

(7.50)

satisfying the Schrödinger–like equation

\[
i \partial_t \langle \text{Action}|\text{Intention}\rangle_{\text{paths}} = H_{\text{Action}} \langle \text{Action}|\text{Intention}\rangle_{\text{paths}},
\]

(7.51)

where \(H_{\text{Action}}\) represents the Hamiltonian (total energy) function available at the state of \(\text{Action}\). Here our ‘golden rule’ is: the higher the \(H_{\text{Action}}\), the lower the microscopic fatigue.
In the connectionist language, our propagator expressions (7.50–7.51) represent activation dynamics, to which our Monitor process gives a kind of backpropagation feedback, a version of the basic supervised learning (7.64).

**Mechanisms of Decision–Making under Uncertainty.** The basic question about our local decision making process, occurring under uncertainty at the intention formation fase $F$, is: Which alternative to choose? (see [Roe et al. (2001); Grossberg (1982); Grossberg (1999); Grossberg (1988); Ashcraft (1994)].) In our path–integral language this reads: Which path (alternative) should be given the highest probability weight $w$? Naturally, this problem is iteratively solved by the learning process (7.43–7.64), controlled by the MONITOR feedback, which we term algorithmic approach.

In addition, here we analyze qualitative mechanics of the local decision making process under uncertainty, as a heuristic approach. This qualitative analysis is based on the micro–level interpretation of the Newtonian–like action $S[x]$, given by (7.45) and figuring both processing speed $\dot{x}$ and LTM (i.e., the force–field $\varphi(x)$, see next subsection). Here we consider three different cases:

1. If the potential $\varphi(x)$ is not very dependent upon position $x(t)$, then the more direct paths contribute the most, as longer paths, with higher mean square velocities $[\dot{x}(t)]^2$ make the exponent more negative (after Wick rotation (7.48)).
2. On the other hand, suppose that $\varphi(x)$ does indeed depend on position $x$. For simplicity, let the potential increase for the larger values of $x$. Then a direct path does not necessarily give the largest contribution to the overall transition probability, because the integrated value of the potential is higher than over another paths.
3. Finally, consider a path that deviates widely from the direct path. Then $\varphi(x)$ decreases over that path, but at the same time the velocity $\dot{x}$ increases. In this case, we expect that the increased velocity $\dot{x}$ would more than compensate for the decreased potential over the path.

Therefore, the most important path (i.e., the path with the highest weight $w$) would be one for which any smaller integrated value of the surrounding field potential $\varphi(x)$ is more than compensated for by an increase in kinetic–like energy $\frac{m}{2} \dot{x}^2$. In principle, this is neither the most direct path, nor the longest path, but rather a middle way between the two. Formally, it is the
path along which the average Lagrangian is minimal,
\[
\langle \frac{m}{2} \dot{x}^2 + \varphi(x) \rangle \rightarrow \min,
\] (7.52)
i.e., the path that requires minimal memory (both LTM and WM, see LSF fields below) and processing speed. This mechanical result is consistent with the ‘filter theory’ of selective attention [Broadbent (1958)], proposed in an attempt to explain a range of the existing experimental results. This theory postulates a low level filter that allows only a limited number of percepts to reach the brain at any time. In this theory, the importance of conscious, directed attention is minimized. The type of attention involving low level filtering corresponds to the concept of early selection [Broadbent (1958)].

Although we termed this ‘heuristic approach’ in the sense that we can instantly feel both the processing speed \( \dot{x} \) and the LTM field \( \varphi(x) \) involved, there is clearly a psycho–physical rule in the background, namely the averaging minimum relation (7.52).

From the decision making point of view, all possible paths (alternatives) represent the consequences of decision making. They are, by default, short–term consequences, as they are modelled in the micro–time–level. However, the path integral formalism allows calculation of the long–term consequences, just by extending the integration time, \( t_{\text{fin}} \rightarrow \infty \). Besides, this averaging decision mechanics – choosing the optimal path – actually performs the ‘averaging lift’ in the LSF: from micro– to the macro–level.

**Force–Fields and Memory in LSF fields**

At the macro–level in the subspace LSF fields we formulate the force–field action principle
\[
\delta S[\varphi] = 0,
\] (7.53)
with the action \( S[\varphi] \) dependent on Lewinian force–fields \( \varphi^i = \varphi^i(x) \) \( (i = 1, ..., N) \), defined as a temporal integral
\[
S[\varphi] = \int_{t_{\text{ini}}}^{t_{\text{fin}}} \mathcal{L}[\varphi] \, dt,
\] (7.54)
with Lagrangian density given by
\[
\mathcal{L}[\varphi] = \int d^n x \, \mathcal{L}(\varphi, \partial_x \varphi^i),
\]
where the integral is taken over all \( n \) coordinates \( x^j = x^j(t) \) of the LSF, and \( \partial_{x^j} \phi^i \) are partial derivatives of the field variables over coordinates.

On the micro–level in the subspace \( LSF_{\text{fields}} \) we have the Feynman–type sum over fields \( \phi^i \) \( (i = 1, \ldots, N) \) given by the adaptive path integral

\[
\langle \text{Action} | \text{Intention} \rangle_{\text{fields}} = \oint \mathcal{D}[w \phi] e^{iS[\phi]} \Rightarrow \text{Wick} \oint \mathcal{D}[w \phi] e^{-S[\phi]}, \quad (7.55)
\]

with action \( S[\phi] \) given by temporal integral \( (7.54) \). (Choosing special forms of the force–field action \( S[\phi] \) in \( (7.55) \) defines micro–level MD & CD, in the LSF fields space, that is similar to standard quantum–field equations, see e.g., [Ivancevic and Ivancevic (2006b)].) The corresponding partition function has the form similar to \( (7.49) \), but with field energy levels.

Regarding topology of the force fields, we have in place \( n \)–categorical Lagrangian–field structure on the Riemannian LSF manifold \( M \),

\[
\Phi^i : [0, 1] \rightarrow M, \Phi^i : \Phi^i_0 \mapsto \Phi^i_1,
\]

generalized from the recursive homotopy dynamics [Ivancevic and Ivancevic (2006b)], using

\[
\frac{d}{dt} f_{x^i} = f_{x^i} \rightarrow \partial_{\mu} \left( \frac{\partial \mathcal{L}}{\partial \Phi^i_{\mu}} \right) = \frac{\partial \mathcal{L}}{\partial \Phi^i},
\]

with \( [x_0, x_1] \mapsto [\Phi^i_0, \Phi^i_1] \).

**Relationship between Memory and Force–Fields.** As already mentioned, the subspace \( LSF_{\text{fields}} \) is related to our memory storage [Ashcraft (1994)]. Its global macro–level represents the long–term memory (LTM), defined by the least action principle \( (7.53) \), related to cognitive economy in the model of semantic memory [Ratcliff (1978); Collins et al. (2001)]. Its local micro–level represents working memory (WM), a limited–capacity ‘bottleneck’ defined by the adaptive path integral \( (7.55) \). According to our formalism, each of Miller’s 7 ± 2 units [Miller (1956)] of the local WM are adaptively stored and averaged to give the global LTM capacity (similar to the physical notion of potential). This averaging memory lift, from WM to LTM represents retroactive interference, while the opposite direction, given by the path integral \( (7.55) \) itself, represents proactive interference. Both retroactive and proactive interferences are examples of the impact of cognitive contexts on memory. Motivational contexts can exert their influence, too. For example, a reduction in task–related recall following the completion of the task is one of the clearest examples of force–field
influences on memory: the amount of details remembered of a task declines as the force–field tension to complete the task is reduced by actually completing it.

Once defined, the global LTM potential $\varphi = \varphi(x)$ is then affecting the locomotion transition paths through the path action principle (7.45), as well as general learning (7.43) and decision making process (7.52).

On the other hand, the two levels of $LSF_{fields}$ fit nicely into the two levels of processing framework, as presented by [Craik and Lockhart (1972)], as an alternative to theories of separate stages for sensory, working and long–term memory. According to the levels of processing framework, stimulus information is processed at multiple levels simultaneously depending upon its characteristics. In this framework, our macro–level memory field, defined by the fields action principle (7.53), corresponds to the shallow memory, while our micro–level memory field, defined by the adaptive path integral (7.55), corresponds to the deep memory.

**Geometries, Topologies and Noise in $LSF_{geom}$**

On the macro–level in the subspace $LSF_{geom}$ representing an $n$–dimensional smooth manifold $M$ with the global Riemannian metric tensor $g_{ij}$, we formulate the geometrical action principle

$$\delta S[g_{ij}] = 0,$$

where $S = S[g_{ij}]$ is the $n$–dimensional geodesic action on $M$,

$$S[g_{ij}] = \int d^n x \sqrt{g_{ij} dx^i dx^j}.$$  \hfill (7.56)

The corresponding Euler–Lagrangian equation gives the geodesic equation of the shortest path in the manifold $M$,

$$\dddot{x}^i + \Gamma^i_{jk} \dot{x}^j \dot{x}^k = 0,$$

where the symbol $\Gamma^i_{jk}$ denotes the so–called affine connection which is the source of curvature, which is geometrical description for noise (see [Ingber (1997); Ingber (1998)]). The higher the local curvatures of the LSF–manifold $M$, the greater the noise in the life space. This noise is the source of our micro–level fluctuations. It can be internal or external; in both cases it curves our micro–LSF.
Otherwise, if instead we choose an \( n \)-dimensional Hilbert–like action (see Misner et al. (1973)),
\[
S[g_{ij}] = \int d^n x \sqrt{\det |g_{ij}|} R,
\]
where \( R \) is the scalar curvature (derived from \( \Gamma^i_{jk} \)), we get the \( n \)-dimensional Einstein–like equation:
\[
G_{ij} = 8\pi T_{ij},
\]
where \( G_{ij} \) is the Einstein–like tensor representing geometry of the LSF manifold \( M \) (\( G_{ij} \) is the trace–reversed Ricci tensor \( R_{ij} \), which is itself the trace of the Riemann curvature tensor of the manifold \( M \)), while \( T_{ij} \) is the \( n \)-dimensional stress–energy–momentum tensor. This equation explicitly states that psycho–physics of the LSF is proportional to its geometry. \( T_{ij} \) is important quantity, representing motivational energy, geometry–imposed stress and momentum of (loco)motion. As before, we have our ‘golden rule’: the greater the \( T_{ij} \)–components, the higher the speed of cognitive processes and the lower the macroscopic fatigue.

The choice between the geodesic action (7.56) and the Hilbert action (7.57) depends on our interpretation of time. If time is not included in the LSF manifold \( M \) (non–relativistic approach) then we choose the geodesic action. If time is included in the LSF manifold \( M \) (making it a relativistic–like \( n \)–dimensional space–time) then the Hilbert action is preferred. The first approach is more related to the information processing and the working memory. The later, space–time approach can be related to the long–term memory: we usually recall events closely associated with the times of their happening.

On the micro–level in the subspace \( \text{LSF}_{\text{geom}} \) we have the adaptive sum over geometries, represented by the path integral over all local (regional) Riemannian metrics \( g_{ij} = g_{ij}(x) \) varying from point to point on \( M \) (modulo diffeomorphisms),
\[
\langle \text{Action}|\text{Intention} \rangle_{\text{geom}} = \int \mathcal{D}[w_{g_{ij}}] e^{iS[g_{ij}]} = \text{Wick} \int \mathcal{D}[w_{g_{ij}}] e^{-S[g_{ij}]},
\]
where \( \mathcal{D}[g_{ij}] \) is diffeomorphism equivalence class of \( g_{ij}(x) \in M \).

To include the topological structure (e.g., a number of holes) in \( M \), we can extend (7.58) as
\[
\langle \text{Action}|\text{Intention} \rangle_{\text{geom/top}} = \sum_{\text{topol.}} \int \mathcal{D}[w_{g_{ij}}] e^{iS[g_{ij}]},
\]
(7.59)
where the topological sum is taken over all connectedness–components of $M$ determined by the Euler characteristic $\chi$ of $M$. This type of integral defines the theory of fluctuating geometries, a propagator between $(n-1)$–dimensional boundaries of the $n$–dimensional manifold $M$. One has to contribute a meaning to the integration over geometries. A key ingredient in doing so is to approximate (using simplicial approximation and Regge calculus [Misner et al. (1973)]) in a natural way the smooth structures of the manifold $M$ by piecewise linear structures (mostly using topological simplices $\Delta$). In this way, after the Wick–rotation (7.48), the integral (7.58–7.59) becomes a simple statistical system, given by partition function 

$$Z = \sum_{\Delta} C_\Delta e^{-S_\Delta}$$

where the summation is over all triangulations $\Delta$ of the manifold $M$, while $C_\Delta$ is the order of the automorphism group of the performed triangulation.

**Micro–Level Geometry: the source of noise and stress in LSF.**
The subspace $LSF_{geom}$ is the source of noise, fluctuations and obstacles, as well as psycho–physical stress. Its micro–level is adaptive, reflecting the human ability to efficiently act within the noisy environment and under the stress conditions. By averaging it produces smooth geometry of certain curvature, which is at the same time the smooth psycho–physics. This macro–level geometry directly affects the memory fields and indirectly affects the (loco)motion transition paths.

**The Mental Force Law.** As an effective summary of this section, we state that the psychodynamic transition functor $T_A : INTENTION_{init} \rightarrow ACTION_{fin}$, defined by the generic path integral (7.70), can be interpreted as a mental force law, analogous to our musculo–skeletal covariant force law, $F_i = mg_{ij}a_j$, and its associated covariant force functor $F_\ast : TT^*M \rightarrow TTM$ [Ivancevic and Ivancevic (2005)].

### 7.1.9 Joint Action Psychodynamics

Cognitive neuroscience investigations, including fMRI studies of human co-action, suggest that cognitive and neural processes supporting co-action include joint attention, action observation, task sharing, and action coordination [Fogassi et al. (2005); Knoblich and Jordan03; Sebanz et al. (2006)]. For example, when two actors are given a joint control task (e.g., tracking a moving target on screen) and potentially conflicting controls (e.g., one person in charge acceleration, the other – deceleration), their joint performance depends on how well they can anticipate each other’s actions. In particular, better coordination is achieved when individuals receive real-time feedback
about the timing of each other’s actions [Sebanz et al. (2006)].

To model the dynamics of this joint action, we associate each of the actors with an \( n \)−dimensional (\( nD \), for short) Riemannian Life-Space manifold, that is a set of their own time dependent trajectories, \( M_\alpha = \{ x^i(t_i) \} \) and \( M_\beta = \{ y^j(t_j) \} \), respectively. Following [Ivancevic and Aidman (2007)], we use the modelling machinery consisting of the joint psycho–physical action (7.67) and the corresponding adaptive path integral (7.70), visualized by the Feynman–like 1–loop diagram.\(^{36}\)

Recently [Ivancevic and Aidman (2007)] we have suggested a generalized motivational/cognitive action, generating Lewinian force–fields [Lewin (1951); Lewin (1997)] on smooth manifolds, extending and adapting classical Wheeler–Feynman action–at–a–distance electrodynamics [Wheeler and Feynman (1949)]. Applying this approach to human co–action, we propose a two–term joint action:

\[
A[x, y; t_i, t_j] = \frac{1}{2} \int_{t_i}^{t_j} \int_{t_i}^{t_j} \alpha_i \beta_j \delta(I_{ij}^2) \dot{x}^i(t_i) \dot{y}^j(t_j) \, dt_i dt_j \\
+ \frac{1}{2} \int_{t_i}^{t_j} g_{ij} \dddot{x}^i(t) \dddot{y}^j(t) \, dt
\]

with \( I_{ij}^2 = [x^i(t_i) - y^j(t_j)]^2 \), (7.60)

where \( IN \leq t_i, t_j \leq OUT \), while \( \dot{x}^i(t_i) = dx^i/dt_i \) and \( \dot{y}^j(t_j) = dy^j/dt_j \) are the corresponding \( nD \) (loco)motion velocities.

The first term in (7.67) represents potential energy between the cognitive/motivational interaction of the two agents \( \alpha_i \) and \( \beta_j \). It is a double integral over a delta function of the square of interval \( I_{ij}^2 \) between two points on the paths in their Life–Spaces; thus, interaction occurs only when this interval, representing the motivational cognitive distance between the two agents, vanishes. Note that the cognitive (loco)motions of the two agents \( \alpha_i[x^i(t_i)] \) and \( \beta_j[y^j(t_j)] \), generally occur at different times \( t_i \) and \( t_j \) unless \( t_i = t_j \), when cognitive synchronization occurs.

The second term in (7.67) represents kinetic energy of the physical interaction. Namely, when the cognitive synchronization in the first term actually takes place, the second term of physical kinetic energy is activated in the common manifold, which is one of the agents’ Life Spaces, say \( M_\alpha = \{ x^i(t_i) \} \).

\(^{36}\)The work presented in this subsection has been developed in collaboration with Dr Eugene Aidman and Mr Leong Yen, both Senior Research Scientists, Land Operations Division, Defence Science & Technology Organisation, Australia.
The adaptive path integral (see [Ivancevic and Aidman (2007)]) represents an infinite–dimensional neural network, corresponding to the psycho–physical action (7.67), reads

$$\langle \text{OUT} | \text{IN} \rangle := \int_{\Sigma} D[w, x, y] e^{iA[x, y; t_i, t]} , \quad (7.61)$$

where the Lebesgue integration is performed over all continuous paths $x^i = x^i(t_i)$ and $y^j = y^j(t_j)$, while summation is performed over all associated discrete Markov fluctuations and jumps. The symbolic differential in the path integral (7.70) represents an adaptive path measure, defined as a weighted product

$$D[w, x, y] = \lim_{N \to \infty} \prod_{s=1}^{N} w^s_{ij} dx^i dy^j , \quad (i, j = 1, ..., n). \quad (7.62)$$

The adaptive path integral (7.70) incorporates the local Bernstein’s adaptation process [Bernstein (1947)] [Bernstein (1935)] according to Bernstein’s discriminator concept desired state $SW(t+1) = current state IW(t) + adjustment step \Delta W(t)$. The robustness of biological motor control systems in handling excess degrees of freedom has been attributed to a combination of tight hierarchical central planning and multiple levels of sensory feedback-based self-regulation that are relatively autonomous in their operation [Bernstein et al. (1996)]. These two processes are connected through a top-down process of action script delegation and a bottom-up emergency escalation mechanisms. There is a complex interplay between the continuous sensory feedback and motion/action planning to achieve effective operation in uncertain environments (in movement on uneven terrain cluttered with obstacles, for example).

Complementing Bernstein’s motor control principles is Brooks’ concept of computational subsumption architectures [Brooks (1985)] [Brooks (1990)], which provides a method for structuring reactive systems from the bottom up using layered sets of behaviors. Each layer implements a particular goal of the agent, which subsumes that of the underlying layers.

For example, a robot’s lowest layer could be “avoid an object”, on top of it would be the layer “wander around”, which in turn lies under “explore the world”. The top layer in such a case could represent the ultimate goal of “creating a map”. This way, the lowest layers can work as fast–responding mechanisms (i.e., reflexes), while the higher layers control the
main direction to be taken in order to achieve a more abstract goal.

The substrate for this architecture comprises a network of finite state machines augmented with timing elements. The subsumption compiler compiles *augmented finite state machine* (AFSM) descriptions into a special-purpose scheduler to simulate parallelism and a set of finite state machine simulation routines. Their networked behavior can be described conceptually as:

$$\text{final state } w(t+1) = \text{current state } w(t) + \text{adjustment behavior } f(\Delta w(t)).$$

The Bernstein weights, or Brooks nodes, $w_{ij}^s = w_{ij}^s(t)$ in (7.71) are updated by the Bernstein loop during the joint transition process, according to one of the two standard neural learning schemes, in which the micro–time level is traversed in discrete steps, i.e., if $t = t_0, t_1, ..., t_s$ then $t + 1 = t_1, t_2, ..., t_{s+1}$:

(1) A self–organized, unsupervised (e.g., Hebbian–like [Hebb (1949)]) learning rule:

$$w_{ij}^s(t+1) = w_{ij}^s(t) + \frac{\sigma}{\eta}(w_{ij}^s_d(t) - w_{ij}^s_a(t)), \quad (7.63)$$

where $\sigma = \sigma(t), \eta = \eta(t)$ denote signal and noise, respectively, while new superscripts $d$ and $a$ denote desired and achieved micro–states, respectively; or

(2) A certain form of a supervised gradient descent learning:

$$w_{ij}^s(t+1) = w_{ij}^s(t) - \eta \nabla J(t), \quad (7.64)$$

where $\eta$ is a small constant, called the step size, or the learning rate, and $\nabla J(t)$ denotes the gradient of the ‘performance hyper–surface’ at the $t$–th iteration.

Both Hebbian and supervised learning are untiled in local decision making processes, e.g., at the intention formation phase (see [Ivancevic and Aidman (2007)]). Overall, the model presents a set of formalisms to represent time–critical aspects of collective performance in tactical teams. Its
applications include hypotheses generation for real and simulation experiments on team performance, both in human teams (e.g., emergency crews) and hybrid human-machine teams (e.g., human-robotic crews). It is of particular value to the latter, as the increasing autonomy of robotic platforms poses non-trivial challenges, not only for the design of their operator interfaces, but also for the design of the teams themselves and their concept of operations.

7.1.10 Feynman–Finsler Approach to Adaptation

Imagine three agents, \(\alpha_i = \alpha_i(t_i)\), \(\beta_j = \beta_j(t_j)\) and \(\gamma_k = \gamma_k(t_k)\), continually evolving and adapting in their own times \(t_i, t_j\) and \(t_k\), performing a goal–driven interaction, that is a joint action of driving a car, so that \(\alpha_i\) does the steering, \(\beta_j\) does the accelerating and \(\gamma_k\) does the braking.

To model this joint agents action, we associate to each of them an \(n\)-dimensional (\(nD\), for short) configuration manifold, that is a set of their own–time dependent trajectories, 

\[
M_{\alpha} = \{x^i(t_i)\}, M_{\beta} = \{y^j(t_j)\} \quad \text{and} \quad M_{\gamma} = \{z^k(t_k)\},
\]

respectively. Their associated tangent bundles contain their individual velocities,

\[
TM_{\alpha} = \{\dot{x}^i(t_i) = dx^i/dt_i\}, TM_{\beta} = \{\dot{y}^j(t_j) = dy^j/dt_j\} \quad \text{and} \quad TM_{\gamma} = \{\dot{z}^k(t_k) = dz^k/dt_k\}.
\]

The joint action happens in the common 3\(n\)D Finsler manifold 

\[
M_J = M_{\alpha} \cup M_{\beta} \cup M_{\gamma},
\]

parameterized by the local joint coordinates dependent on the common time \(t\). That is, 

\[
M_J = \{q^r(t), r = 1, \ldots, 3n\}.
\]

Geometry of the joint manifold \(M_J\) is defined by the Finsler metric function 

\[
ds = F(q^r, dq^r),
\]

and the Finsler tensor 

\[
C_{rst}(q, \dot{q}),
\]

defined by (see \[Rund (1959)\] and \[Ivancevic and Ivancevic (2007b)\])

\[
C_{rst}(q, \dot{q}) = \frac{1}{4} \frac{\partial^2 F^2(q, \dot{q})}{\partial q^r \partial q^s} = \frac{1}{2} \frac{\partial g_{rs}}{\partial \dot{q}^r \partial \dot{q}^s},
\]

From the Finsler definitions \((7.65)\)–\((7.66)\), it follows that the partial interaction manifolds, \(M_{\alpha} \cup M_{\beta}, M_{\beta} \cup M_{\gamma}\) and \(M_{\alpha} \cup M_{\gamma}\) have Riemannian structures with the corresponding interaction kinetic energies, 

\[
T_{\alpha\beta} = \frac{1}{2} g_{ij} \dot{x}^i \dot{y}^j, \quad T_{\alpha\gamma} = \frac{1}{2} g_{ik} \dot{x}^i \dot{z}^k \quad \text{and} \quad T_{\beta\gamma} = \frac{1}{2} g_{jk} \dot{y}^j \dot{z}^k.
\]

Now, following \[Ivancevic and Aidman (2007)\], we use the modelling machinery consisting of:
1. Adaptive joint action (7.67)–(7.69) at the top-master level, describing the externally-appearing deterministic, continuous and smooth dynamics, and

2. Corresponding adaptive path integral (7.70) at the bottom-slave level, describing a wildly fluctuating dynamics including both continuous trajectories and Markov chains.

At the master level, the adaptive joint action reads

\[ A[t_i, t_j, t_k; t] = \frac{1}{2} \int_{t_i}^{t_j} \int_{t_k}^{t_j} \alpha_i(t_i) \beta_j(t_j) \gamma_k(t_k) \delta(I_{ijk}^2) x^i(t_i) y^j(t_j) z^k(t_k) dt_i dt_j dt_k \]  

\[ + \frac{1}{2} \int_t W_{rs}^M(t, q, \dot{q}) \dot{q}^r \dot{q}^s dt \quad \text{(where } IN \leq t_i, t_j, t_k, t \leq OUT) \]  

with

\[ I_{ijk}^2 = [x^i(t_i) - y^i(t_j)]^2 + [y^j(t_j) - z^k(t_k)]^2 + [z^k(t_k) - x^i(t_i)]^2, \]  

The first term (7.67) in the joint action, contains the cognitive intention Lagrangian of the three agents coming into the joint action. It is a triple integral over their own timescales. The actual physical action (given by the second term (7.68)) would happen only if their timescales synchronize, that is in the case \( t_i = t_j = t_k \). Otherwise, (7.67) is just the sum of their individual kinetic potentials. The sub-term \( \delta(I_{ijk}^2) \), given by (7.69) is the delta function of their mutual “cognitive distance” \( I_{ijk}^2 \) that vanishes upon the synchronization.

The second action term (7.68) represents adaptive kinetic energy of their physical interaction. Namely, when the previous “cognitive synchronization” in the first term actually occurs, the second term of physical kinetic energy is activated in the common time \( t \). Then we have the joint physical motions of the three agents, \( \alpha_i, \beta_j \) and \( \gamma_k \), physically moving in the joint Finsler coordinates \( \{q^r(t), r = 1, ..., 3n\} \), along the common timescale \( t \), in their joint 3nD manifold \( M_J \). This joint kinetic energy is adaptive, represented by “master joint synaptic weights” \( W_{rs}^M(t, q, \dot{q}) \), which is the Riemannian metric tensor (7.65) allowed to evolve in time.

At the slave level, the adaptive path integral (see Ivancevic and Aidman (2007)), representing an infinite-dimensional neural network, corresponding to the adaptive joint action (7.67), reads

\[ \langle OUT | IN \rangle := \oint D[w; x, y, z; q] e^{iA[t_i, t_j, t_k; t]}, \]  

\[ (7.70) \]
where the Lebesgue integration is performed over all continuous paths $x^i = x^i(t_i), y^j = y^j(t_j), z^k = z^k(t_k)$ and $q^r = q^r(t)$ while the summation is performed over all associated discrete Markov fluctuations and jumps. The symbolic differential in the path integral (7.70) represents an adaptive path measure, defined as a weighted product

$$D[w; x, y, z; q] = \lim_{N \to \infty} \prod_{S=1}^{N} \frac{1}{w_{ijkr}^S} dx^i dy^j dz^k dq^r,$$

$$i, j, k = 1, ..., n; \quad r = 1, ..., 3n.$$  

(7.71)

The “slave synaptic weights” $w_{ijkr}^S = w_{ijkr}^S(t)$ in (7.71) are updated according to a self-organized, unsupervised (e.g., Hebbian-like [Hebb (1949)]) learning rule (7.63), or a certain form of a supervised gradient descent learning (7.64), which are both naturally used for the local decision making process occurring at the intention formation phase (see [Ivancevic and Aidman (2007)]. In the cognitive psychology framework, our adaptive path integral (7.70) can be interpreted as semantic integration (see [Bransford and Franks (1971); Ashcraft (1994)]).


Bell, J.S: On the Einstein Podolsky Rosen Paradox. Physics 1, 195, (1964)


Bellman, R., Kalaba, R.: Selected papers on mathematical trends in control theory, Dover, New York, (1964)


Quantum Leap

Blackmore, D.L., Prykarpatsky, Y.A., Samulyak, R.V.: The Integrability of Lie-
Bousso, R., Polchinski, J.: Quantization of Four Form Fluxes and Dynamical Neutralization of the Cosmological Constant. JHEP 0006, 006, (2000)
Quantum Leap


Chen, X., Tian, G.: Ricci flow on Kahler-Einstein surfaces. Invent. Math. 147,
Bibliography

Colless, M., Dalton, G., Maddox, S., Sutherland, W.: The 2dF Galaxy Redshift
Quantum Leap

Bibliography

DePietri, R.: On the relation between the connection and the loop representation of quantum gravity, Class. and Quantum Grav. 14, 53–69, (1997)


Dirac, P.A.M.: Relativity Quantum Mechanics with an Application to Compton


Esposito, G.: Quantum cosmology from three different perspectives. arXiv:gr-


Freeman, W.J.: Tutorial on neurobiology: from single neurons to brain chaos.
Freeman, W.J.: Random activity at the microscopic neural level in cortex sustains
is regulated by low dimensional dynamics of macroscopic cortical activity.
Friberg, O.: A Set of Parameters for Finite Rotations and Translations. Comp.
Friedman, J.L., Papastamatiou, N.J., Simon, J.Z.: Unitarity of Interacting Fields
Frittelli, S., Kozameh, C., Newman, E.T., Rovelli, C., Tate, R.S.: Fuzzy space–
time points from the null–surface formulation of general relativity, Class.
Quantum Gravity. 14, A143, (1997)
Fulling, S.: Aspects of quantum field theory in curved space–time. Cambridge
B146, 90, (1978)
Gérard, R., Tahara, H.: Singular nonlinear partial differential equations, Aspects
Gade, P.M., Hu, C.K.: Synchronous chaos in coupled map lattices with small-world
Garcia, P.: Connections and 1-jet fibre bundles. Rend. Sem. Univ. Padova, 47,
227, (1972)
Garcia, P.: Gauge algebras, curvature and symplectic structure. J. Diff. Geom.,
12, 209, (1977)
Gardiner, C.W.: Handbook of Stochastic Methods for Physics, Chemistry and
works and identifying compound mode of action via expression profiling.
Giachetta, G., Mangiarotti, L.: Gauge-invariant and covariant operators in gauge
Giachetta, G., Sardanashvily, G.: Stress-energy–momentum of affine–metric grav-
Giachetta, G., Sardanashvily, G.: Dirac Equation in Gauge and Affine-Metric
Giachetta, G., Mangiarotti, L., Sardanashvily, G.: New Lagrangian and Hamilton-
Giachetta, G., Mangiarotti, L., Sardanashvily, G.: Covariant Hamilton equations
Giachetta, G., Mangiarotti, L., Sardanashvily, G.: Covariant geometric quanti-
Giachetta, G., Mangiarotti, L., Sardanashvily, G.: Geometric quantization of
Giachetta, G., Mangiarotti, L., Sardanashvily, G.: Lagrangian supersymmetries
Gibbons, G.W.: Aspects of Supergravity Theories. In Supersymmetry, Super-
gravity, and Related Topics, eds. F. del Aguila, J.A. de Azcarraga and L.E.
Gibbons, G.W., Hartle, J.B.: Real Tunneling Geometries and the Large-scale
Gibbons, G.W., Hartle, J.B.: Real Tunneling Geometries and the Large-Scale
Gibbons, G.W., Hawking, S.W.: Action integrals and partition functions in quan-
Gibbons, G.W., Hawking, S.W., Perry, M.J.: Path Integrals and the Indeﬁnity
Gilks, W.R., Richardson, S., Spiegelhalter, D.J.: Markov Chain Monte Carlo in
Gisin, N., Ribordy, G., Tittel, W., Zbinden, H.: Quantum cryptography Rev.
Glauber, R.J.: The Quantum Theory of Optical Coherence. Phys. Rev. 130,
2529-2539, (1963)
Quantum Leap


Hawking, S.W., Ellis, G.F.R.: The large scale structure of space-time. Cambridge Univ. Press, (1973)
Bibliography

Quantum Leap

Ivancevic, V.: Some Possibilities of Multilayered Neural Networks Application in Biomechanics of Muscular Contractions, Human Motion and Sport Train-
Bibliography


Quantum Leap


Leibniz, G.W.: Monadology, (1714)


Lychagin, V.V.: Contact geometry and nonlinear second order differential equations. Uspekhi Mat. Nauk, 34, 137-165 (in Russian), (1979)
Meyer, K.R., Hall, G.R.: Introduction to Hamiltonian Dynamical Systems and
213–234, (1979)


Nanopoulos, D.V.: Theory of Brain Function, Quantum Mechanics and Super-
863, (2005)
Quantum Leap

Bibliography

Reshetikhin, N., Takhtajan, L.: Deformation Quantization of Kalher Manifolds.
Quantum Leap

Russell, B.: Mysticism, Logic and Other Essays, London: Longmans, Green, (1918)
Sardanashvily, G.: The Lyapunov stability of first order dynamic equations with
   Rev. Lett. 65, 22622265, (1990)
Satarić, M.V., Tuszyński, J.A., Zakula, R.B.: Kinklike excitations as an energy-
   402, 469, (1993)
Schlichenmaier, M.: Berezin–Toeplitz Quantization and Berezin’s Symbols for Arbitrary Compact Kähler Manifolds, in Coherent States, Quantization
   and Gravity, M. Schlichenmaier et al. (eds.) Polish Scientific Publishers
   PWN, Warsaw, (2001)
Schmaltz, T.: Nicolas Malebranche, The Stanford Encyclopedia of Philosophy,
   Cambridge, MA, (1994)
Shen, Z.: Riemann–Finsler geometry with applications to information geometry, unpublished material, (2005)
Spinoza, B.: Tractatus Theologico–Politicus (A Theologico-Political Treatise), (1670)
Steenrod, N.: The Topology of Fibre Bundles, Princeton Univ. Press, Princeton,
Tucker, A.W.: unpublished material (1950)


Wiesenfeld, K., Colet, P., Strogatz, S.H.: Synchronization Transitions in a Dis-
Wiesenfeld, K., Colet, P., Strogatz, S.H.: Frequency locking in Josephson arrays:
Wiggins, S.: Introduction to Applied Dynamical Systems and Chaos. Springer,
New York, (1990)
Williams, R.M.: Recent progress in Regge calculus. Nucl. Phys. B 57, 73–81,
(1997)
Williams, R.M., Tuckey, P.: Regge Calculus: A bibliography and brief review.
Class. Quant. Grav. 9, 1409, (1992)
Willmore, J., Howinson, S.: Option Pricing: Mathematical Models
Wilson, D.: Nonlinear Control, Advanced Control Course (Student Version),
Wineland, D.J., Bollinger, J.J., Itano, W.M., Moore, F.L.: Spin squeezing and
reduced quantum noise in spectroscopy. Phys. Rev. A 46, R6797–R6800,
(1992)
Winfree, A.T.: Biological rhythms and the behavior of populations of coupled
Witten, E.: Supersymmetry and Morse theory. J. Diff. Geom., 17, 661–692,
(1982)
291, (1986)
Witten, E.: 2+1 Gravity as an Exactly Soluble Model. Nucl. Phys. B311, 46-78,
(1988)
(1988)
Witten, E.: Space-Time and Topological Orbifolds. Phys. Rev. Lett. 61, 670–673,
(1988)
Witten, E.: On the structure of the topological phase of two–dimensional gravity.
Witten, E.: Introduction To Cohomological Field Theories. Int. J. Mod. Phys. A
6, 2775, (1991)
Witten, E.: Mirror manifolds and topological field theory. In Essays on mirror
Witten, E.: Chern–Simons gauge theory as a string theory. Prog. Math. 133,
Witten, E: String Theory Dynamics in Various Dimensions. Nucl. Phys. B 443,
This page intentionally left blank
Index

abductive reasoning, 17
ability to think abstractly, 12
absorption and emission operators, 46
achieved, 508, 758
action, 64, 120, 128
action functional, 64
action–amplitude picture, 67, 738
action–at–a–distance
electrodynamics, 766
adaptive path integral, 749
adaptive path measure, 507, 757, 761
AdS–Schwarzschild solution, 451
affine, 516
affine connection, 514
affine–connection, 63
all, 536, 549
American option, 242
amplitude, 75, 337
analytic philosophy, 651
analytical philosophy, 647
anthropic principle, 304
anthropic reasoning, 410, 417
anthropic string landscape, 367
anti de Sitter space, 125
anticommutation relations, 60
any, 370
array of N Josephson Junctions, 601
arrow of time, 701
Ashtekar connection, 291
Ashtekar phase–space formalism, 291
Atiyah axioms, 276
axiomatic system, 648
Banach space, 105
Batalin–Vilkovisky quantization, 149
Battle of the Sexes, 527
behaviorism, 654
Bekenstein–Hawking entropy, 222
Bekenstein–Hawking formula, 267
Bell test experiments, 666
Bell’s inequality, 586
Bell’s theorem, 586, 658
best response, 527
Betti numbers, 261
bi–partite quantum system, 543
Big–Bang, 268
Big–Crunch/Big–Bang transition, 411
binary system, 645
biological images, 507
Birkhoff’s Theorem, 435
black hole dynamics, 220
Black–Scholes–Merton formula, 232
Bogomol’nyi monopoles, 149
Born–Dirac rule, 676
Bose–Einstein condensate, 76
Bose–Einstein condensation, 503, 539
Bose–Einstein statistics, 14
Bose–Fermi super-multiplets, 66
bosonic string theory, 177
bottom–up cosmology, 508
Index

Cooper pairs, 596
cooperative game theory, 523
coordination game, 527
Copenhagen interpretation, 657
correlation function, 123
correlation functions, 149
correspondence principle, 179
cosmology, 342
covariant derivative, 517
covariant Dirac equation, 61
covariant force functor, 755
covariant force law, 755
covariant muscular forces, 499
creativity, 11
criterion of falsifiability, 643
critical current, 599, 600
critical rationalism, 642
critical slowing down, 105
cumulative distribution function, 69
current-phase relation, 603
curvature, 514
cuspy patterns, 256

Dalambertian wave operator, 58
dark radiation, 439
de Sitter space, 307, 370
de Sitter universe, 371
de Sitter vacua, 109
de Sitter vacuum, 370
decohere, 355
decohere, 25
Decoherence functional, 325, 359
decoherence functional, 326, 355
definition of a Nash equilibrium, 354
desired, 508, 758
Deutsch algorithm, 504
diactics, 652
diffeomorphism invariance, 290
different, 565
diffusion equation, 122, 123
dilaton field, 178
Dirac, 60
Dirac adjoint spinor, 60
Dirac equation, 60
Dirac interaction picture, 60
Dirac operator, 60, 278

Dirac quantum commutator, 313
Dirac rules for quantization, 23
Dirac wave spinor, 60
Dirac’s electrodynamic action
principle, 40
Dirac’s formalism, 20
Dirichlet branes, 179
discrete characteristic projector, 32
discrete eigenvalues, 32
discrete Laplacian, 616
discrete sine–Gordon equation, 614
discrete spectral form, 32
discrete spectrum, 32
dissipation, 706
dissipative structures, 67
distribution function, 68
dominant strategy, 549
dominant strategy equilibrium, 526
Donaldson polynomials, 149
Donaldson theory, 133, 149
double–slit experiment, 302
dualism, 711
dynamical chaos, 256
dynamical laws, 309

Eccles model of synaptic activation, 514
economics, 521
effective theory, 363
Ehrenfest theorem, 513
Einstein, 63
Einstein equation, 109, 309, 332, 171
Einstein equations, 64
Einstein manifold, 370
Einstein’s field equation, 371
Einstein’s laws of radiation, 19
Einstein–Hilbert action, 63, 110, 274
439
electric current, 500
electrical current, 508
electrical field, 508
electrical muscular stimulation, 601
electrical potential, 509
electrodynamic field action principle, 448
electrodynamic flux tubes, 510
### Index

<table>
<thead>
<tr>
<th>Term</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friedmann equation</td>
<td>307</td>
</tr>
<tr>
<td>Friedmann’s equation</td>
<td>428</td>
</tr>
<tr>
<td>frozen accidents</td>
<td>312</td>
</tr>
<tr>
<td>frustrated XY models</td>
<td>608</td>
</tr>
<tr>
<td>Fubini’s Theorem</td>
<td>470</td>
</tr>
<tr>
<td>function</td>
<td>21</td>
</tr>
<tr>
<td>function collapse</td>
<td>660</td>
</tr>
<tr>
<td>functional electrical stimulation</td>
<td>499</td>
</tr>
<tr>
<td>functional view</td>
<td>637</td>
</tr>
<tr>
<td>fundamental laws</td>
<td>309</td>
</tr>
<tr>
<td>fundamental theory</td>
<td>363</td>
</tr>
<tr>
<td>Gödel’s incompleteness theorem</td>
<td>648</td>
</tr>
<tr>
<td>game matrix</td>
<td>526</td>
</tr>
<tr>
<td>game situations</td>
<td>572</td>
</tr>
<tr>
<td>game theory</td>
<td>529</td>
</tr>
<tr>
<td>game tree</td>
<td>523</td>
</tr>
<tr>
<td>games</td>
<td>524</td>
</tr>
<tr>
<td>gamma–EEG waves</td>
<td>671</td>
</tr>
<tr>
<td>gates</td>
<td>692</td>
</tr>
<tr>
<td>gauge condition</td>
<td>66</td>
</tr>
<tr>
<td>gauge theory</td>
<td>510</td>
</tr>
<tr>
<td>gauge–invariant phase difference</td>
<td>622</td>
</tr>
<tr>
<td>Gauss–Bonnet formula</td>
<td>261</td>
</tr>
<tr>
<td>Gauss–Codazzi equation</td>
<td>447</td>
</tr>
<tr>
<td>Gauss–Kronecker curvature</td>
<td>264</td>
</tr>
<tr>
<td>Gaussian curvature</td>
<td>115</td>
</tr>
<tr>
<td>general relativity</td>
<td>263</td>
</tr>
<tr>
<td>332</td>
<td></td>
</tr>
<tr>
<td>generalized quantum theory</td>
<td>326</td>
</tr>
<tr>
<td>generating functional</td>
<td>319</td>
</tr>
<tr>
<td>generic energy condition</td>
<td>333</td>
</tr>
<tr>
<td>genus</td>
<td>261</td>
</tr>
<tr>
<td>geodesic action</td>
<td>514</td>
</tr>
<tr>
<td>geodesic deviation</td>
<td>517</td>
</tr>
<tr>
<td>geodesic equation</td>
<td>514</td>
</tr>
<tr>
<td>geodesics</td>
<td>63</td>
</tr>
<tr>
<td>geometric action principle</td>
<td>516</td>
</tr>
<tr>
<td>geometrical consistency relation</td>
<td>464</td>
</tr>
<tr>
<td>geometrodynamical functor</td>
<td>740</td>
</tr>
<tr>
<td>ghost number</td>
<td>180</td>
</tr>
<tr>
<td>Glaever tunneling junction</td>
<td>599</td>
</tr>
<tr>
<td>Gibbons–Hawking boundary</td>
<td>473</td>
</tr>
<tr>
<td>Gibbs ensemble</td>
<td>50</td>
</tr>
<tr>
<td>Giti</td>
<td>14</td>
</tr>
<tr>
<td>global space–time hyperbolicity</td>
<td>331</td>
</tr>
<tr>
<td>globally–coupled Kuramoto model</td>
<td>619</td>
</tr>
<tr>
<td>goal–driven interaction</td>
<td>759</td>
</tr>
<tr>
<td>Goldstone theorem</td>
<td>699</td>
</tr>
<tr>
<td>Gottesman–Knill theorem</td>
<td>593</td>
</tr>
<tr>
<td>governing equilibrium dynamics</td>
<td>523</td>
</tr>
<tr>
<td>Green’s function</td>
<td>122</td>
</tr>
<tr>
<td>Green’s functions</td>
<td>512</td>
</tr>
<tr>
<td>Greisen–Zatsepin–Kuzmin limit</td>
<td>373</td>
</tr>
<tr>
<td>groups</td>
<td>573</td>
</tr>
<tr>
<td>growth–rate functions</td>
<td>539</td>
</tr>
<tr>
<td>Hadamard gate</td>
<td>593</td>
</tr>
<tr>
<td>halting problem</td>
<td>588</td>
</tr>
<tr>
<td>Hamilton–Jacobi equation</td>
<td>22</td>
</tr>
<tr>
<td>Hamiltonian</td>
<td>311</td>
</tr>
<tr>
<td>314</td>
<td>969</td>
</tr>
<tr>
<td>harmonic oscillator</td>
<td>519</td>
</tr>
<tr>
<td>Hawking’s Euclidean quantum gravity</td>
<td>275</td>
</tr>
<tr>
<td>Hawking’s no–boundary wave function</td>
<td>524</td>
</tr>
<tr>
<td>Hawking’s no–boundary wave</td>
<td></td>
</tr>
<tr>
<td>function of the universe</td>
<td>321</td>
</tr>
<tr>
<td>Hawking–Moss (HM) instanton</td>
<td>413</td>
</tr>
<tr>
<td>Hawking–Moss instanton</td>
<td>372</td>
</tr>
<tr>
<td>Hawking–Turok instanton</td>
<td>420</td>
</tr>
<tr>
<td>Heisenberg</td>
<td>19</td>
</tr>
<tr>
<td>Heisenberg picture</td>
<td>30</td>
</tr>
<tr>
<td>309</td>
<td>313</td>
</tr>
<tr>
<td>313</td>
<td></td>
</tr>
<tr>
<td>Heisenberg uncertainty principle</td>
<td>22</td>
</tr>
<tr>
<td>301</td>
<td>656</td>
</tr>
<tr>
<td>656</td>
<td>662</td>
</tr>
<tr>
<td>662</td>
<td>686</td>
</tr>
<tr>
<td>Heisenberg uncertainty relation</td>
<td>29</td>
</tr>
<tr>
<td>Heisenberg’s method</td>
<td>19</td>
</tr>
<tr>
<td>her</td>
<td>558</td>
</tr>
<tr>
<td>Hermitian (self–adjoint) linear operator</td>
<td>22</td>
</tr>
<tr>
<td>Hermitian inner (scalar) product</td>
<td>8</td>
</tr>
<tr>
<td>Hermitian inner product</td>
<td>9</td>
</tr>
<tr>
<td>Hermiticity</td>
<td>327</td>
</tr>
<tr>
<td>359</td>
<td></td>
</tr>
<tr>
<td>heuristic quantization rule</td>
<td>294</td>
</tr>
<tr>
<td>hidden variable theory</td>
<td>662</td>
</tr>
<tr>
<td>hierarchy problem</td>
<td>430</td>
</tr>
<tr>
<td>Hilbert basis</td>
<td>16</td>
</tr>
<tr>
<td>Hilbert space</td>
<td>11</td>
</tr>
<tr>
<td>109</td>
<td>314</td>
</tr>
<tr>
<td>116</td>
<td></td>
</tr>
<tr>
<td>Hodgkin–Huxley neural</td>
<td>514</td>
</tr>
<tr>
<td>holographic hypothesis</td>
<td>698</td>
</tr>
<tr>
<td>holonomic brain model</td>
<td>838</td>
</tr>
</tbody>
</table>
homeostatic neuro–muscular feedbacks, 507
horospherical coordinates, 471
Hubble parameter, 307, 419
Hubble volume, 307
Human Biodynamics Engine, 512
human mind, 11
ideal, 564
imaginary time, 125
imagination, 635
imitative, 539
Immirzi parameter, 273
incompatible observables, 660
incompleteness of description, 22
incompleteness of quantum mechanics, 656
indeterminacy principle, 670, 672
infinite–dimensional Hilbert space, 33
infinite–dimensional neural network, 739
inflation, 306
inflaton field, 307
initial condition, 310
initial conditions, 309
initial point, 305
inner product space, 9
integrable mechanical systems, 510
intellect, 635
intellegentia, 11
intelligence, 11, 641
intelligible world, 14
interactionism, 642
interior, 536
intertwining tensor, 296
intuitionistic logic, 653
Ito lemma, 233
Ito stochastic integral, 71
Jacobi equation, 517
Jacobi fields, 517
Jacobi flows, 517
joint action, 739
Josephson constant, 589
Josephson current, 589
Josephson current–phase relation, 600
Josephson effect, 596
Josephson interferometer, 596
Josephson junction, 595
Josephson junction ladder, 608
Josephson tunnelling, 596, 599
Josephson voltage–phase relation, 600
Josephson–junction quantum computer, 598
jump, 136
K–theory, 170
Kähler gauge, 220
Kähler potential, 222
Kaluz–Klein action, 470
Kaluz–Klein manifold, 471
Kaluz–Klein theory, 469
Kaluz–Klein tower, 470
Killing vector, 279
kinetic energy of muscular contraction, 711
kinetic–energy estimator, 197
kink–phason resonance, 672
Kleyn–Gordon equation, 41
Kleyn–Gordon Lagrangian, 96
Kortveg–De Vries equation, 513
Kraus operators, 552
Kuramoto order parameter, 629
ladder, 620
Lagrangian, 93
Lagrangian density, 90, 505, 744
Lagrangian–field structure, 506
Landau gauge, 615
language, 11
Large Hadron Collider, 129
late time, 457
law of contradiction, 655
law of the excluded middle, 655
learning, 12
learning rate, 758
least action principle, 504, 505
Lebesgue integral, 33
Lewinian force–field theory, 739
Lie bracket, 54
Lie super–algebra, 135
Index

line element, 510
linear operator equation, 122
linear representation, 23
linear superposition, 2
linearized Hamiltonian dynamics, 257
Liouville equation, 72
local Bernstein’s adaptation process, 757
locality, 660
locally–coupled Kuramoto model, 619
locally–optimal solution, 710
logicism, 648
long–range connections, 630
long–range correlation, 699
long–range order, 597
loop algebra, 289
loop quantum gravity, 285
loop representation, 300
loop variables, 270
Lorentz equation of motion, 57
Lorentz force, 509
Lorentz transformations, 41
Lorentz–invariant theories, 172
Lorentzian de Sitter space, 419
Lorentzian dynamical triangulations, 1172
Lorentzian manifold, 370
Lorentzian region, 410
Lorentzian–de Sitter metric, 346
Lorenz gauge, 65
low energy effective action, 149
low–energy–limit, 457
Lyapunov exponent, 256
Mach–Zehnder interferometer, 7
macro–level averaging lift, 511
macroscopic force–velocity model, 513
magnetic field, 508
Maldacena duality, 431
many–worlds theories, 504
map from strategy profiles to pay–offs, 523
Markov assumption, 72
Markov chain, 69
Markov stochastic process, 69
mass–action interpretation, 536
mass–shell, 66
Master equation, 72
Matching Pennies, 528
Mathai–Quillen formula, 145
matrices, 60
Maurer–Cartan connection, 147
Maxwell, 63
Maxwell electrodynamics, 508
Maxwell stress–energy tensor, 471
McCumber parameter, 625
mean, 16
mean–field theory, 259
Measure of Interference, 355
mechanical action, 511
meditation, 723
memory, 635
mental abilities, 11
mental force law, 755
metastable sector, 367
method of least squares, 16
metric, 415
metric manifold, 288
Metropolis algorithm, 103
microscopic sliding filament model, 513
microtubules, 667
mind, 635
mind–body problem, 640
Minkowski action, 100
Minkowski metric, 429
Minu, 14
mixed strategies, 556
mixed strategy, 528
moduli, 443
moduli field, 457
moduli space of supersymmetric vacua, 374
molecular soliton model of muscular contraction, 572
momentum conservation law, 463
momentum constraint equation, 345
monetary, 527
monism, 646
Quantum Leap

Monte Carlo method, 101
Montonen–Olive duality, 149
Moore’s law, 590
Morse function, 258, 262
Morse lemma, 267
Morse numbers, 262
Morse theory, 128, 258, 262, 316
moves, 524
multi–scalar–tensor theory, 459
multicanonical sampling, 107
multigrid method, 106
muscle–fat manifold, 520
muscular contraction action principle, 511
muscular Riemannian configuration manifold, 505
muscular Riemannian metrics, 511
myofibrillar action propagators, 512
Nambu–Goto action, 176
Nash equilibrium, 527, 536, 549
Nash’s program, 532
nature, 524
negative tension brane, 456
Neumann boundary condition, 181
neural path integral, 248
neurobiology, 654
Newman–Penrose equation, 332
Newman–Penrose formalism, 332
Newton’s laws of motion, 309
Newtonian determinism, 309
Newtonian equation of motion, 311
no boundary proposal, 172
no–boundary initial condition, 310
no–boundary proposal, 306
no–boundary quantum state, 412
no–boundary quantum state of the universe, 364
no–boundary wave function, 364
noncommutative space, 170
nonlinear control theory, 179
nonlinear Schrödinger equation, 513
nonlinear sigma model, 176
nonlocal behavior, 658
nonlocal process, 657
norm, 5
normal, 517, 524, 525
normalization, 327, 350
normalization condition, 29
normalized state, 5
normalized state of the universe, 364
not, 457
Occam’s razor, 304
Ockham’s razor, 429
on–shell reducible, 156
open string theories, 171
optimal answer, 555
optimal policy, 758
optimization, 645
orbifold, 170
orbit Hilbert space, 38
order parameters, 369
oriented strings, 172
orthogonal sum, 34
orthogonality, 8
overall probability amplitude, 506
overdamped junction, 621
overdamped ladder, 631, 633
overdamped limit, 604
parallel transport, 63
parallelogram law, 10
Pareto optimal, 550
Pareto optimality, 530
partition function, 98, 255
path integral, 67, 227, 338
path–integral expression, 94
path–integral formalism, 91
path–integral formulation, 90
path–integral quantization, 90
Pauli, 57
Pauli gate, 593
pay–off, 524
Index

penny flipover, 576
Penrose paradox, 668
perception, 639, 732
perceptual world, 14
perfect & imperfect Nash equilibria, 533
personality, 11
perturbative expansion methods, 507
perturbative path integral, 100
perturbative string theory, 177
perturbative vacuum, 367
Peyrard–Bishop system, 258
phase, 75
phase coherence, 507
phase difference, 600, 616
phase transition, 253, 258, 368
phase transitions, 105
phase-space path integral, 86
phenomenology, 653
physically–controllable systems of nonlinear oscillators, 620
plan, 11
Planck brane, 430
Planck constant, 505
Planck length, 171
Planck's constant, 4, 22, 303
Poincare recurrence time, 380
Poincare recurrences, 380
point at infinity, 3
Poisson bracket, 58
Poisson detection statistics, 76
Poisson evolution equation, 513
polarization law, 9
political philosophy, 637
Polyakov action, 176
Pontryagin class, 151
Pontryagin Maximum Principle, 227
Ponzano–Regge ansatz, 274
Ponzano–Regge quantization, 270
population–statistic, 534
positive chirality condition, 152
Positivity, 527, 359
Prakrti, 641
pre–Big–Bang cosmologies, 411
Principle of Democracy, 441
principle of locality, 650
Principle of relativity, 40
Principle of superposition, 559
Principle of superposition of states, 22
Prisoner’s Dilemma, 625
Prisoners’ Dilemma, 542
pristine, 620
probabilistic transition micro–dynamics, 506
probability, 16, 21
probability amplitude, 90
probability amplitudes, 25
probability conservation law, 22
probability distribution function, 624
probability–density current, 61
product manifold, 153
propagator, 99
psychology, 11
pull–back, 470
pulse, 730
pure continuous spectral form, 36
pure discrete spectral form, 36
pure strategy, 556
Purusha, 641
quantization, 23
quantum algorithm, 594
quantum brain, 738
quantum chromodynamics, 429
quantum coherent state, 75
quantum computation, 670
quantum computers, 588, 596
quantum consciousness, 729
quantum cosmology, 394
quantum cryptography, 663, 685, 587
quantum decoherence, 661
quantum electrodynamics, 510
quantum entanglement, 25, 668, 660
quantum evolution equation, 513
quantum field theory, 465, 167
quantum gravity, 108
quantum Hamilton’s equations, 20
quantum Hamiltonian operator, 513
quantum histories, 314
quantum information, 385
Quantum mechanics of closed systems, 353
quantum Monty Hall, 576
quantum observables, 600
quantum operator, 299
quantum optics, 586
quantum parallelism, 594
quantum pictures, 30
quantum simulator, 589
quantum state ket–vector, 2
quantum strategy, 548
quantum superposition, 24
quantum teleportation, 586
quantum truel, 576
quantum–mechanical wave function, 596
qubits, 588, 591, 596
radion, 469
random, 628
random variable, 68
random walk, 69
Rarita–Schwinger equations, 66
rationalistic, 534
Ray–Singer torsion, 149
Raychaudhuri’s equation, 435
reason, 11
recalled, 538
recursive homotopy dynamics, 752
Regge calculus, 111, 359
Regge geometries, 119
Regge simplicial action, 116
regularities in time, 309
reheating, 307
reinforcement learning, 758
relative acceleration, 514
relativistic Hamiltonian form, 40
Relativistic Heavy Ion Collider, 373
replicator dynamics, 539
resistive loading, 601
resistively & capacitively–shunted junction, 622
resistively–shunted Josephson junctions, 615
resistively–shunted junction, 621
resonance, 598
Ricci tensor, 370, 336
Riemann curvature tensor, 370, 516
Riemann sphere, 3
Riemannian metric tensor, 504
Riesz Representation Theorem, 10
rigged Hilbert space, 33
risk dominance, 536
rules of the game, 543
Russell Paradox, 648
S–matrix, 409
Sachs–Wolfe formula, 493
saddle point approximation, 345
Samadhi, 723
Sankhya school, 641
scalar curvature, 518
scalar electrical potential, 508
scattering, 434
Scholastic tradition, 445
Schrödinger’s wave, 21
Schrödinger, 20
Schrödinger equation, 2, 70, 85, 311
Schrödinger equation, 4
Schrödinger evolution, 674
Schrödinger picture, 2, 30, 99
Schrödinger’s Cat, 660
Schrödinger’s cat paradox, 304
Schrödinger’s method, 20
Schwarz–type TQFT, 161
Schwarzschild metric form, 337
Schwinger formalism, 511
search for truth, 643
second quantization, 46
second–order phase transitions, 368
self–organized, 506, 768, 761
semantic integration, 761
semantic theory of truth, 643
semi–classical history, 411
semiclassical approximation, 369
semiclassical effective action, 369
set of instructions, 524
set of probability coefficients, 20
shape operator, 202
shortest path, 514
Index

signal, 508, 758
simple statistical system, 521
simplicial approximation, 521
single valued in a time, 328
Singularity Theorems, 302
slow–roll inflation, 372
small world, 622
small–world networks, 620
social situations, 523
solid state physics, 585
solitary model of muscular excitation–contraction, 510, 511
soliton, 510
solution in maxi-min-strategies, 550
source, 126
sources and sinks, 67
space, 508
space of all weighted paths, 512
space–time coarse graining, 328
spectral theory, 19
spin connection, 291
spin connection 1–form, 153
spin foam models, 281
spin network, 296
spin network theory, 297
spin networks, 110
spin–wave modes, 614
spinor notation, 332
spontaneous symmetry breaking, 369
spring constant, 119
staging algorithm, 106
standard action, 93
Standard Model, 265, 285, 409
Standard model, 68
statistical error, 103
step size, 508, 758
stimulus–response-type, 501
stochastic integral, 70
strategic, 529
strategic form, 524
strategic interaction, 523
strategies, 543
strategy, 524
strategy profile, 525
stream of photons, 5
string corrections, 174
string cosmology, 269
string landscape, 409
string tension, 176
string theory landscape, 308
string–field, 219
string–field–theory action, 180
strong Church–Turing thesis, 589
strong energy condition, 333
Sturm–Liouville problem, 123
substantial view, 636
subsumption architectures, 757
sum over fields, 510
sum over fractal geometries, 520
sum over histories, 512
sum–over–fields, 364
sum–over–geometries, 364
sum–over–histories, 502
super–field, 210
super–space, 210
superconducting–normal transition, 608
superconductivity, 589, 596, 720
superfluidity, 720
supergravity, 65
supermoduli–space, 374
superposition, 510
superstring theory, 177
supersymmetric Yang–Mills theory, 177
supersymmetry, 177
supervised gradient descent learning, 758, 761
Suppose \( f \) meets [IN]. If a population state \((s, p)\) is stationary under the associated dynamics, 538
Suppose \( f \) meets [IN]. If a population state \( s \) is stationary under the associated dynamics, 538
Suppose \( g \) meets [POS], and consider the associated population dynamics, 540
syllogism, 14
symmetry, 122
symmetry–breaking transition, 368
synaptic weights, 507
synchronization of coupled nonlinear oscillators, 619
systematic error, 102
tachyon field, 178
tangent dynamics equation, 257
tangential, 517
Tao, 646
techincolor models, 428
tensor perturbations, 470
TeV–brane, 430
The Last Problem, 528
the outcome that maximizes the product of the players’ utilities, 531
The Principle of Superposition, 327
theory of everything, 308, 309
theory of fluctuating geometries, 520
there exists a unique solution, 531
thermodynamic partition function, 338
thought, 635
thought experiment, 655
time average, 504
time–dependent Schrödinger equation, 24
time–ordered product, 122
time–ordering operator, 606
time–reversal symmetry, 368
top–down cosmology, 306
topological Bogom“lnyi action, 155
topological hypothesis, 254
topological invariant, 151, 152
topological quantum field theory, 128
topological Theorem, 256
total action, 504
total Hilbert state–space, 54
total spectral form, 52
total spectral measure, 53
trace anomaly inflation, 409
transformation rules, 659
transition, 23
transition amplitude, 674
transition entropy, 512
transition functor, 402
transition probability, 70
transition probability distribution, 239
Trotter decomposition formula, 105
truth as correspondence, 643
truth–in–itself, 653
tunnelling, 367
tunnelling rate, 370
turiya state, 722
twisted Dirac equation, 154
two forms of quantum electrodynamics, 40
two interpretations, 634
two player game, 543
two–person zero–sum strategic game, 545
two–player quantum game, 548
two–point function, 124
type–I gauge theories, 64
type–II gauge theories, 65
umbrella sampling, 107
underdamped junction, 621
underdamped ladder array, 631
unitary evolution, 2
universal Turing machine, 588
unstable resonance, 385
unsupervised, 507
utility, 527
utility functionals, 548
utility functions, 549
vacuum distribution, 44
vacuum state, 70
virial estimator, 107
virtual paths, 67
volatility, 233
voltage–to–frequency converter, 600
volume–weighted probability distribution, 423
W–cycle, 106
warp factor, 130
warped geometry, 128
warped product metric, 428
wave equation, 22
Index

wave function, 311
wave function of the universe, 303
wave–function reduction, 304
wave–particle duality, 75
weak energy condition, 332
weakly–connected neural network, 639
weights, 758
well–defined input–output function, 592
Weyl scalars, 332
Weyl spinor, 152
Weyl tensor, 332, 430
Weyl–fluid, 338
Wheeler–DeWitt equation, 345
when players move, 524
where \( f \) is defined in, 540
whirling modes, 614
whirling regime, 617
Wick rotation, 119, 183, 337, 507
Wiener process, 72
Wigner function, 701
Wigner’s friend paradox, 661
will, 635
Wilson loop, 149
Winfree–type phase models, 619
wisdom, 11
Witten, 150
Witten’s TQFT, 128
world–sheet, 170
Yang–Baxter equation, 133
Yang–Lee Theorem, 254
Yang–Mills gauge theory, 129
Yang–Mills theories, 469
Yang–Mills theory, 270, 474
Zamolodchikov metric, 215
zero sum game: one player’s loss is another player’s gain, 528
zero–sum game, 528