Title
Asymptotic Analysis of Spin Networks with Applications to Quantum Gravity

Author
Haggard, Hal Mayi

Publication Date
2011

Peer reviewed|Thesis/dissertation

Asymptotic Analysis of Spin Networks with Applications to Quantum Gravity

by

Hal Mayi Haggard

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Robert G. Littlejohn, Chair

Professor Ori J. Ganor

Professor Alan D. Weinstein

Spring 2011

Asymptotic Analysis of Spin Networks with Applications to Quantum Gravity

Hal Mayi Haggard

Abstract

Asymptotic Analysis of Spin Networks with Applications to Quantum Gravity

by

Hal Mayi Haggard

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Robert G. Littlejohn, Chair

This work initiates a study of the semiclassical limit of quantum gravity using a geometrical formulation of WKB theory and the Hamilton-Jacobi equation. Few conceptual principles are available to guide physicists in the construction of a quantum theory of gravity. Experimentally accessible signals are notoriously difficult to extract from existing proposals and one of the few reasonable constraints that we can impose is that the proposals agree with general relativity in the classical limit. Because general relativity is such a rich classical theory this is a non-trivial condition, one that has yet to be quantitatively achieved by any theory of quantum gravity. The main focus of the dissertation is on the semiclassics of $SU(2)$ spin networks. Spin networks play an important role in the loop gravity approach to quantum gravity, where they furnish a convenient and geometrically meaningful basis for the Hilbert space. Previous work on the semiclassics and asymptotics of spin networks have focused on a coherent state approach. Here we provide alternative methods based on geometrical Lagrangian manifolds. This new perspective is complementary; for example, calculation of amplitudes is very straightforward, and should open new research avenues.

The thesis consists of two parts. In the first part, Foundations, we review the geometrical formulation of WKB theory and introduce the theory of spin networks from the beginning. These chapters make the tools and applications covered in this thesis readily accessible to new researchers and open the door to further cross-fertilization between researchers in semiclassics and loop gravity. In the second part, Applications, we focus on two applications of semiclassical theory to objects arising in loop gravity. In the loop approach to quantum gravity the geometry of space becomes discretized. Our first application is a derivation of the semiclassical spectrum and wavefunctions of the volume operator of a tetrahedral grain of space. A comparison of this spectrum with that found in loop gravity shows excellent agreement. This provides a simplified derivation of the quantization of space that strengthens earlier proposals along these lines. The second application is an asymptotic formula for the $9j$-symbol including its amplitude, phase, and all of the phase adjustments. The $9j$-symbol is a more complex spin network than has been treated at this level of detail before and

arises as part of the vertex amplitude in spinfoams, the loop gravity analog of the path integral approach to quantum gravity. More broadly this quantitative result provides further motivation for developing the asymptotics of higher $3nj$-symbols; in the long term these asymptotics, which are accurate even for small quantum numbers, may furnish an effective computational tool for bridging loop gravity predictions to testable experiments.

In loving memory of my father.

To my mother, whose irrepressible joy and wonderment I aspire to share with all those around me.

Contents

List of Figures

List of Tables

Acknowledgments

My heart is full of more gratitude to all of you than I could possibly express. These words will not suffice.

For so many years of guidance and support I thank my Berkeley advisor Robert Littlejohn. Robert, your willingness to always start at the beginning and your astonishing ability to see into the very core of concepts will always guide me. I am grateful for the time we have spent together. To my advisors in Pavia, Mauro Carfora and Annalisa Marzuoli, also go a very special thanks. Your immense warmth and generosity has been a great joy to me. You have so graciously opened the door to Italy and Europe and have both made me feel at home and given me the opportunity to flourish. To my committee members, Ori Ganor and Alan Weinstein, thank you both for your patience and your enthusiasm. I look forward to many future discussions of the subtleties of mathematics and quantum gravity.

Eugenio Bianchi, I can not imagine a more delightful collaborator. You are curious, lively, intelligent and so full of warmth. I hope that there will be many more opportunities for us to explore the mysteries of Nature together.

My visits to Europe have been supported by the kind generosity of many friends and collaborators. I hope to one day repay your generosity tenfold and my home is always open to you. Thank you to: Paola Abenante, Vincenzo Aquilanti, Giorgio Immirzi, Sarah Kliegman, Elena Magliaro, Antonino Marcianò, Claudio Perini, Matteo Polettini, Carlo Rovelli, Matteo Smerlak, Simone Speziale, Francesca Vidotto. Sincere apologies to any that I have failed to mention; know that it is simply the vicissitudes of memory.

A special thanks goes to Philipp Rostalski, Bernd Sturmfels and Cynthia Vinzant for discussions of algebraic geometry and the solution of systems of polynomial equations.

For their financial support I would like to acknowledge the Center for Pure and Applied Mathematics at Berkeley, the France-Berkeley Fund, A. Carl Helmholtz, the Phi Beta Kappa Northern California Association and the UC Berkeley Dissertation-Year Fellowship program. Another thank you to Robert for supporting my funding applications. I have also received a number of travel grants through the National Science Foundation; these grants were available due to the hard work of Jorge Pullin and Don Marolf in applying for them. Thank you!

This dissertation would not have been possible without you my dear family. You have given me love and confidence, kindness and courage—everything. You are the very fabric of my life. I love you.

It seems that my family is an ever widening circle. Great gratitude to those who have crafted my home here in Berkeley. Kater, Becky, West, Molly, Stefan, Marty, Amber, Sunny, Emily, may I always live in the glow of your love and care. Angie, Badr, Benji, Joel, I would not have made it through my experience at Berkeley

without you. A greater demonstration of courage has never been mine to participate in. May the Compass Project still flourish when we are all old and gray. Andrew, Nadir and Liang you have been much more than colleagues; I hope that we continue to learn and celebrate together.

Crystal and Paige there is too much to say to you. But then, that is always how I feel with you. Thank you for always listening.

My family at Yoga Mandala you have given me more than I ever expected or knew I wanted. Thank you for expanding my horizons.

Hrimati and little Boo thank you for your incredible patience and strength. Few authors owe such a debt to their children before they are even born but this has been our special circumstance. I love you both!

Chapter 1

Introduction

To a large extent this work is inspired by the beauty of classical mechanics. Classical mechanics is not fundamental, it has been supplanted by the quantum, but neither is it outdated; it provides invaluable insight into our understanding of quantum theory. This we aim to demonstrate in the course of this dissertation.

This introduction should naturally begin explaining the theory of quantum gravity, outlining its successes and domains of validity and discussing outstanding problems. The fact that we cannot do this yet, despite more than fifty years of effort, is immensely frustrating to physicists. This frustration also represents an opportunity. Indeed, one view of Einstein's magnificent approach to doing physics is that he was able to sit in the midst of cognitive dissonance, to hold two opposing ideas together long enough, for remarkable higher levels of insight to arise. The questions that arise in quantum gravity are worth this investment. What are space and time? What is spacetime? How does matter come to shape spacetime and how does spacetime come to move matter? No amount of ink put to page can capture the subtlety and richness of these questions. Try them out again, explain time to a loved one, a colleague or even to yourself and see how much room there is to grow into.

Throughout our lives we constantly build models: models of social interactions and networks, models of healthy eating and exercise, models of spatial relations and navigation. These models are like linguistic analogies; they elegantly represent aspects of reality and simultaneously introduce anomalies, that is, failures and limitations of the model to reflect reality. The greatest joy of a scientist is to become aware of anomalies in our understanding of nature and to devote our efforts to more deeply understanding nature's resolution of these apparent inconsistencies. The great mismatch between our descriptions of nature using general relativity and those using quantum mechanics represents an anomaly worthy of our strongest efforts and deepest patience.

Progress towards a theory of quantum gravity has not been monolithic; instead there are a number of approaches: string theory, causal dynamical triangulations, causal sets and loop quantum gravity to name just a few. This thesis focuses on approaches inspired by loop quantum gravity. This does not represent an overarching

critique of other approaches but rather was dictated largely by circumstance. The realization that lying behind spin networks there was a collection of classically integrable systems was our original inspiration for studying loop gravity in detail. However, the loop gravity approach has many insights that recommend it and the connection of spin networks to analytically controllable calculations is certainly amongst these.

It would be impractical to give a comprehensive overview of loop gravity here. There are several excellent books and reviews that serve this purpose [1, 2, 3, 4, 5, 6]. However, to motivate and conceptually situate our discussion of spin networks it is important to survey the role they play in the theory. For this purpose we steal the elegant exposition of Rovelli [7]. The idea is to orient the theory with respect to the more familiar quantum field theories, namely quantum electrodynamics (QED) and quantum chromodynamics (QCD). This approach is somewhat axiomatic and only intended to give a conceptual framework not to explain the origin of the loop quantization. In this manner we avoid discussion of the subtleties involved in the quantization of constrained systems and issues of regularization. These subtleties are important, essential even, to a complete understanding of the theory but on a first pass they are difficult to appreciate. At any rate, the books and reviews cited above may be consulted for these details. We will also restrict attention to the Hamiltonian version of the theory. The Hamiltonian quantization was the first to be attempted and is the theory properly termed loop quantum gravity. This approach involves a splitting of spacetime into space and time, an unappealing move for a generally covariant theory. By contrast much recent work has focused on a covariant, path integral approach called spin foams. When we want to refer to these two approaches collectively we will use the shortened modifier "loop gravity."

Our restriction to Hamiltonian loop quantum gravity here is for the purposes of theoretical and computational simplicity. This will allow us to focus on the gauge group $SU(2)$ and the theory of spin networks is well developed and geometrically clearer here. Much progress has also been made in the covariant context both in the Euclidean $SO(4)$ theory and in Lorentzian approaches; this work will be touched on in the historical review of semiclassics below.

Before defining the Hilbert space of loop quantum gravity we introduce notation and conventions for graphs. An abstract graph is a set of L links labelled by ℓ, N nodes labelled by n and an association of each link to the two nodes it connects. For a directed graph, Γ, the association of the nodes to a link is oriented and there are two maps s and t with s taking the link ℓ to its source node, $s : \ell \mapsto s(\ell)$, and t taking ℓ to its target node, $t : \ell \mapsto t(\ell)$. An $SU(2)$ spin network is a directed graph with an additional label j on each link that associates the spin-j irreducible representation of $SU(2)$ to that link and these additional labels are referred to as a coloring of the graph. Figure 1.1 provides a schematic of a simple graph.

The Hilbert space of loop quantum gravity is built on the graph Hilbert space $\tilde{\mathcal{H}}_\Gamma$. This is the Hilbert space of an $SU(2)$ gauge theory living on the graph Γ, $\tilde{\mathcal{H}}_\Gamma \equiv L^2(SU(2)^L/SU(2)^N)$. Elements of this space, $\psi \in \tilde{\mathcal{H}}_\Gamma$, are functions of an

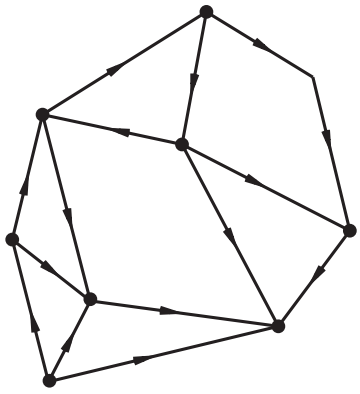


Figure 1.1: Schematic diagram of a simple graph.

$SU(2)$ group element h_ℓ on each link, $\psi = \psi(h_\ell)$, that are invariant under gauge transformations

$$\psi(h_\ell) \mapsto \psi(g_{s(\ell)} h_\ell g_{t(\ell)}^{-1}) \qquad g_n \in SU(2) \tag{1.1}$$

at the nodes adjoining the link.

A graph $\Gamma^{'}$ is a subgraph of Γ if there is a map taking the nodes and links of $\Gamma^{'}$ to a subset of those of Γ in a manner that preserves the orientation of $\Gamma^{'}$. This gives rise to an inclusion of $\tilde{\mathcal{H}}_{\Gamma'} \subset \tilde{\mathcal{H}}_\Gamma$ where $\psi \in \tilde{\mathcal{H}}_\Gamma$ is in the subspace $\tilde{\mathcal{H}}_{\Gamma'}$ if it only depends on $h_{\ell'}$ with $\ell^{'}$ a link of Γ that is the image of a link of $\Gamma^{'}$. The importance of these inclusions is that they give rise to an equivalence relation between states: two states ψ_1 and ψ_2 are equivalent if they depend on graphs Γ_1 and Γ_2 that can be identified with one another as subgraphs of Γ or if Γ_1 and Γ_2 can be identified with one another by an automorphism.

Let $\mathcal{H}_\Gamma \equiv \tilde{\mathcal{H}}_\Gamma / \sim$ where $\sim$ is the equivalence relation just defined. The Hilbert space of loop quantum gravity is

$$\mathcal{H} = \lim_{\Gamma \to \infty} \mathcal{H}_\Gamma. \tag{1.2}$$

The limit $\Gamma \to \infty$ makes sense because of the subgraph inclusion: For any two graphs Γ_1 and Γ_2 we can always find a graph containing them both and extend consideration to the functions on this larger graph. More formally this is made precise by the mathematical notion of a projective or inverse limit.

In practice one always works with a finite truncation of the degrees of freedom of the theory and so our interest will center on $\mathcal{H}_\Gamma$. This is analogous to cutting off the degrees of freedom of the electric field by restricting attention to the interaction of only a finite number of quanta, say N particles, instead of working with the full Fock space. There are a number of parallels between this construction and those performed in QED and QCD. In QED the bare N particle Hilbert space $\tilde{\mathcal{H}}_N$ is moded out by permutations of the particles to give the symmetrical states of $\mathcal{H}_N = \tilde{\mathcal{H}}_N / \sim$.

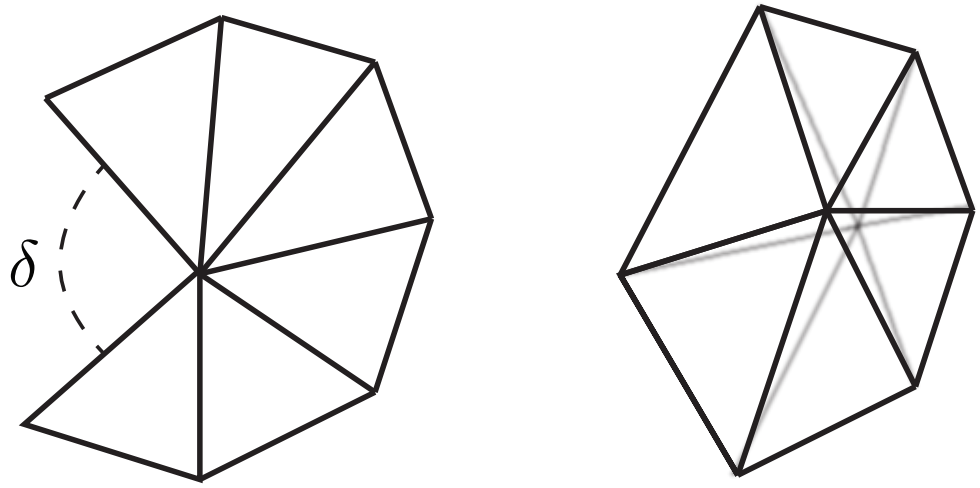
The Hilbert space $\mathcal{H}_N$ supports at most N particles but interactions can change the number of excited quanta. Similarly, the subgraph inclusion allows for interactions to change the number of links and nodes determining the state $\psi(h_\ell)$. On the other hand $\tilde{\mathcal{H}}_\Gamma$ is precisely the sort of space used in lattice gauge theories. A lattice gauge theory with lattice Γ and gauge group G has Hilbert space $\mathcal{H}_{\Gamma G} = L^2\left(G^L/G^N\right)$ and gauge transformations act in an analogous manner to (1.1).

There are also important distinctions between loop quantum gravity and these two theories having to do with the physical interpretation of the theory. In general relativity the gravitational field is spacetime and the realization emerges that fields are always measured with respect to other fields. In loop gravity the quantum state supported by the graph Γ is understood as an excitation of N grains or discrete chunks of space, the nodes of the graph. These grains are not located with respect to anything else. Rather, their adjacency is given by the links connecting the nodes. This beautifully captures the relational character of general relativity and at the same time generalizes the role of the lattice in lattice gauge theories as space.

As we already mentioned the restriction to a finite graph represents the truncation of the infinite degrees of freedom of the field to a finite number. Researchers are still working to understand the full picture of what this truncation represents physically; however, in the semiclassical limit a very nice picture emerges. In Chapter 4 of this dissertation we provide a detailed argument showing that in the semiclassical limit an F-valent node of a spin network corresponds to a convex polyhedron with F faces. Moreover, in the case of a 4-valent node, i.e. a tetrahedron, we show that the Bohr-Sommerfeld quantization of the volume of this polyhedron is in excellent agreement with the spectrum of the volume operator acting at this node. Thus in the semiclassical limit the nodes of the graph Γ can be thought of as a collection of polyhedra. The polyhedra are glued along their faces according to whether the corresponding nodes are joined by a link. Note that unless special restrictions are placed on Γ the shapes of the faces are not necessarily congruent (they can even be distinct polygons, e.g. one triangular and one quadrilateral). The very general discrete geometries arising from this gluing have been dubbed twisted geometries (see [8, 9] for many more details).

In the case that the shapes of the faces match and the whole thing glues up into a nice simplicial decomposition, these discrete geometries were introduced into general relativity and studied by Regge, [10, 11]. Regge had a magnificent insight into how the coordinate dependence, which is pure gauge, might be removed from general relativity. He considered a collection of flat Euclidian tetrahedra, flat in the sense of having zero Riemannian curvature on their interiors, and noted that each tetrahedron's geometry was determined by the specification of its six edge lengths.[1] Despite the flatness of these pieces they can be used to approximate a curved manifold. This works in any

[1]There is a discrete ambiguity for some sets of given edge lengths. For example, there are thirty distinct tetrahedra, the maximum possible number, with edge lengths $\{7, 8, 9, 10, 11, 12\}$, see [12]. If necessary, assume that a disambiguation has been specified along with the edge lengths.



(a) A collection of triangles joined along their edges.

(b) Gluing the last edge results in a conical shape that approximates a smoothly curved surface.

Figure 1.2: Regge triangulation: The discrete curvature of this triangulation is concentrated at its central vertex, the bone of this example. The failure of the angles, $\{\alpha_i\}_{i=1,\ldots,n}$, meeting at a bone to add up to 2π is a measure of the curvature, this is the deficit angle $\delta = 2\pi - \sum_i \alpha_i$.

number of dimensions but is easiest to visualize in two dimensions, see Figure 1.2.

In this framework curvature is distributional and concentrated at the $(n-2)$-dimensional intersections of the n-dimensional simplices, these are called bones. Curvature is defined here by considering the rotation effected on a vector after parallel translation around a loop enclosing a bone. This rotation turns out to be perpendicular to the bone and in an amount equal to the deficit angle of the bone (see Fig. 1.2 for the definition of deficit angle).

Because the simplices are flat a frame chosen at any point of the simplex can be extended to the whole interior. However, orientation of this frame is completely arbitrary and this is the $SU(2)$ gauge freedom of the theory. The association between the graph nodes and polyhedra reveals the parallel between this gauge freedom and the gauge freedom of the nodes in the quantum Hilbert space $\mathcal{H}_\Gamma$ described above, (1.1). The group element labeling a link is the parallel transport that takes the frame of one of these simplices to that of the neighboring simplex. For closed-loop holonomies and the appropriate choice of connection (the Ashtekar-Barbero connection), this captures the extrinsic geometry of 3-dimensional space as it sits in the 4-dimensional spacetime. All of this has been introduced to indicate the effect of truncating the in-

finite degrees of freedom of the gravitational field. To summarize, in the semiclassical limit this looks like a piecewise linear approximation of space by polyhedral pieces that approximate curvature by the manner in which they are glued together. It is important to note that this polyhedral picture is only semiclassical; in the quantum regime the node quanta have the usual quantum fuzziness and it doesn't make sense to ascribe a classical geometry to them. Furthermore, generic quantum states are superpositions of the spin network states we have been describing.

Two important issues still need to be discussed: the character of the semiclassical limit mentioned above and the role of time. It is necessary to distinguish the finite graph truncation from the semiclassical limit. The finite graph truncation is not directly a limitation on the size of the spacetime, it can be large or small, rather it is like a mode expansion. If we truncate consideration to the first m modes of a box of length L_{box} then we will only be able to probe distance scales of order L_{box}/m, whether this is a small or large distance depends on the chosen value of L_{box}. Instead the semiclassical limit is one in which the spins j coloring the links of the graph are taken large. It is in this limit that polyhedra can be associated to the nodes of the graph and the size of the polyhedra's face areas and volume grow with j. In fact, in formulating spin networks as basis states of loop gravity, see section 3.5, we will see that the face areas and volume become the quantum numbers of the states. So the semiclassical limit is a large quantum number limit, physically interpreted as a large distance limit and mimicking the old fashioned correspondence principle limits of early quantum theory. The relevant scale for comparison is the Planck area $\ell_{\text{Pl}}^2 = \hbar G$.[2]

There is not yet a satisfactory understanding of the role of time in quantum gravity. For an overview see [13] and the diversity of view points is well represented by the entries of the Foundational Questions Institute's essay contest. One approach, the only one we will discuss here, is to view the spin network states as boundary states ([14]). In quantum field theory we would tend to choose two such states and ask for the transition amplitude between them. However, in a covariant spacetime, without an a priori foliation into constant time slices an alternative is to choose a single boundary state and to view it as completely surrounding the four dimensional region of interest. The dynamics fills in the interior of the boundary and associates a "transition" amplitude to this single boundary state. This leads into the idea of a spin foam; a boundary spin network graph is interpolated with a choice of two complex and an amplitude is calculated for this spin foam. These are like higher dimensional Feynman diagrams and like Feynman diagrams the rules for finding transition amplitudes arise from three simple requirements [15]: locality, superposition and local Lorentz invariance. The spin foam formalism is outside the scope of this dissertation

[2]More precisely, loop quantum gravity has a free parameter γ, the Barbero-Immirzi parameter, and the proper scale is $\ell_{\text{loop}}^2 = 8\pi\gamma\hbar G$. Requiring that loop gravity calculations of black hole entropy agree with Hawking's semiclassical calculation indicate that γ is order unity.

and we once again refer to the reviews at the beginning of this introduction. This qualitative introduction is intended to orient the study of asymptotics of spin networks within the larger context of loop gravity. We turn now to a brief overview of the semiclassical study of spin networks and spin foams.

Research into the semiclassics of spin networks began with the work of Ponzano and Regge on the $6j$-symbol [11]. In this work a deep understanding of WKB theory and of the $6j$-symbol led them to guess a quantitatively accurate asymptotic formula for the symbol. This paper is remarkable: the phase of their formula, S_{PR}, is the action for the boundary contributions of Regge's discrete model for gravity in terms of simplices and can be seen as a discretization of the Gibbons-Hawking-York term of general relativity [16]; it contains several conjectures about the $9j$-symbol that are correct (Chapter 5); and the conclusions include the first construction of a "state sum," or discrete path integral, model for three dimensional gravity.

Proofs and extensions of the $6j$ asymptotics have been an active area of research [17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27]. Studies of the $10j$-symbol, a close relative to the $15j$-symbol, helped to spur the development of new spin foam models at the beginning of the millennium [28, 29, 18, 19]. In recent years the Nottingham group has extensively studied the asymptotics of the $15j$-symbol [30, 31, 32, 33, 34] and relations to the EPRL and FK vertex proposals for spin foams [35, 36]. This work has successfully demonstrated that the phase of the asymptotic EPRL and FK vertices are the Regge action, the simplicial analog of the Einstein-Hilbert action.

Much of this work has focused on coherent states and their asymptotics. These methods have been invaluable for calculating the phase contributions to the asymptotic formulae but explicit computation of amplitudes has remained a challenge in this framework. Explicit formulae for the amplitudes allow the identification of the caustic set and the wavefunction amplitude is generally largest on this set. The $9j$-symbol discussed in Chapter 4 has a rich caustic structure with enhancements of the wave function at points on the interior of the classically allowed region as well as those usually occurring at the boundary between the classically allowed and forbidden regions. The techniques developed here allow calculation of amplitudes for higher $3nj$-symbols. The coherent state formalism can also be cast in terms of (complex) Lagrangian manifolds, the main tools of this thesis, and an interesting avenue for future research would be to develop a combined coherent state and Lagrangian manifold formalism that simultaneously allows easy calculation of phases and amplitudes. This is a very exciting time to be working on quantum gravity; theorists have an unprecedented opportunity to predict a completely new phenomenon. Semiclassics can help to shape a theory that will make such a prediction successful.

The plan of this thesis is as follows: In Chapter 2 we review the geometrical formulation of WKB theory. The heart of the chapter is an introduction to Lagrangian manifolds. The semiclassical approximation of wavefunctions is built around these special subsets of phase space. The general theory is illustrated throughout the chapter by a detailed treatment of the semiclassics of the 2D oscillator. The semiclassical

core is bookended by two sections that describe the geometry of the phase space of the 2D oscillator and connect it to the theory of angular momentum. This connection is Schwinger's model for angular momentum.

The theory of spin networks is introduced in Chapter 3 and the whole formalism is developed from scratch. We begin by illustrating the graphical technology in the familiar context of the vector algebra of $\mathbb{R}^3$ and then introduce invariant theory. Invariant theory provides perspective on the graphical formalism, elucidating both its mechanics and scope. The chapter ends with an introduction to $SU(2)$ spin networks and an explanation of their role as a basis for the Hilbert space of loop quantum gravity.

The style of the first two chapters of this dissertation perhaps requires a brief explanation. These chapters are not, like most dissertation chapters, expositions of traditional research projects. Instead they survey the foundational material that is used in the research that we have been pursuing into the semiclassics of quantum gravity. The time and effort that has gone into their inclusion has been at the cost of including further chapters on the research performed during my tenure at Berkeley (in particular very little discussion of [37] and [24] has been included in this dissertation). This was a conscious choice. The modern day lore, often repeated to dissertation writers in the sciences, is that no one will read your thesis. With this lore in mind I decided to shift focus and write a part of this thesis with an audience of advanced undergraduates and graduate students in mind. These first two chapters are written with the hope of inviting new researchers into this area, they attempt to be self contained and provide perspective. The focus is on drawing connections between many areas so that the reader feels they can stand on ground they know and find their own way into the hinter lands of research. The material discussed in both chapters is spread throughout the literature and it felt as though this slowed my own entrance into these topics and so I have attempted to provide many citations. More advanced topics, such as the detailed discussion of Maslov indices in Chapter 2 and aspects of the treatment of invariant theory in Chapter 3 will hopefully maintain researcher's interest as well.

The quantization of space is taken up in Chapter 4. We introduce the volume operator of loop gravity and use the WKB theory of Chapter 1 to find the semiclassical spectrum and wavefunctions of a tetrahedral grain of space. We explore the limiting values of this spectrum, finding interesting scaling in the small volume limit. The chapter ends with partial extensions of this work to the case of a convex polyhedron with 5 faces, "triangular prisms."

Chapter 5 summarizes our discovery of an asymptotic formula for the $9j$-symbol. We present the formula and discuss its geometrical origins and symmetries. We briefly conclude in Chapter 6 and list a wide variety of problems to be attacked with the tools described in this dissertation.

Part I

Foundations

Chapter 2

Semiclassics and Geometry

This chapter introduces semiclassical theory.[1] The body of the thesis is concerned with the application of semiclassical techniques to blossoming but still speculative proposals for a quantum theory of gravity. The nature of this work is necessarily interdisciplinary; most physicists have only a passing familiarity with semiclassical techniques coming from a short introduction to WKB theory early in their studies or a brief application during a research project. On the other hand, quantum gravity is a rapidly changing field and only experts are abreast of the current themes. This chapter and the next develop the background necessary to understand semiclassical theory and spin networks respectively. Spin networks are the main tools in our applications to quantum gravity. Specifically, this chapter attempts to get the reader up to speed with the application of semiclassical theory rapidly and painlessly. To this end, the theory is mostly discussed in the context of examples and many of the elegant, abstract aspects of the theory are not developed. We aim to pique the readers interest rather than to survey the field formally. For more comprehensive reviews of semiclassics see [41, 42, 38, 43] (the last from a mathematical perspective).

The primary example used to illustrate semiclassical techniques will be the 2D harmonic oscillator. This example is sufficiently simple to transparently illustrate all aspects of its semiclassical treatment. On the other hand, it is sufficiently rich to lead to applications at the forefront of research into the semiclassics of quantum gravity.

The chapter begins with a discussion of the classical evolution of the 2D oscillator and then proceeds to introduce some generalities about semiclassics. Semiclassical theory can be hierarchically organized by the relative difficulty of calculating each piece of the wavefunction of a system. From easiest to hardest these are:

- the amplitude of the wavefunction
- the quantization conditions

[1] The presentation of this chapter is heavily influenced by Robert Littlejohn's notes from a 1988 course on advanced classical mechanics and the paper [38]. Some material on the 2D oscillator is based on [39] and the visualizations draw from [40].

- the phase of the wavefunction
- and the Maslov indices.

Following the brief general introduction to semiclassics, the application of semiclassical techniques is illustrated by exploring each of these computational facets using the 2D oscillator. The last part of this section contains a detailed discussion of Maslov indices, a more advanced topic that the reader may choose to return to later. Finally, the chapter ends with Schwinger's treatment of angular momentum theory and its ties to symmetry and reduction theory (Poisson and symplectic reduction). These topics draw heavily on the 2D oscillator and provide another motivation for its introduction earlier in the chapter. Schwinger's model will form the backbone for our understanding of spin networks and our applications of semiclassics to quantum gravity. The reader familiar with semiclassical techniques is warned that the material in this chapter is classic; the chapter is illustrated with many figures and can be efficiently skimmed by looking over these figures and their captions.

2.1 Classical mechanics of the 2D oscillator

There was a glorious experiment and art exhibit that used to make occasional appearances at the Santa Fe children's museum. It consisted of a rectangular platform suspended from the ceiling by four chains attached at the corners of the platform. The platform was riddled with circular holes each of which could hold a marker. Below the platform was a large table layered with butcher's paper. When you reached the front of the line you selected several markers of various colors and chose which holes to slot them into (half of the initial conditions). Then you gave the platform a push (the other half of the initial conditions) and off it went creating exquisite art.

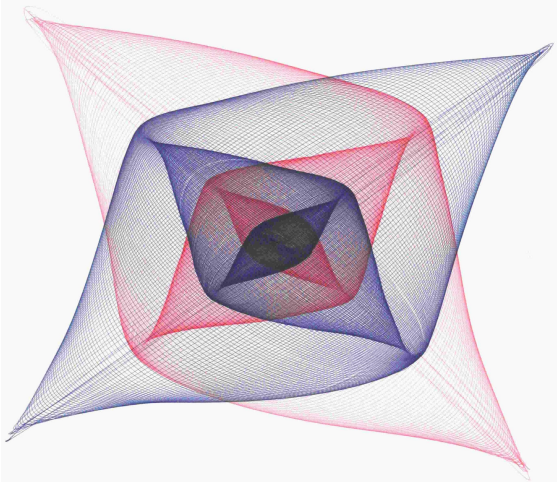
For small amplitude oscillations the platform, called a harmonograph, exhibits two dimensional harmonic oscillations. Physicists call the curves the markers trace out Lissajous figures. Figure 2.1(a) contains an example of a Lissajous figure made using a simple, robust harmonograph crafted by Jonathan Lansey [44]. In the harmonic regime the phase space of the platform is $\mathbb{R}^4$ with coordinates (x_1, p_1, x_2, p_2). The motion can be described using the Hamiltonian

$$H = \frac{1}{2}(p_1^2 + p_2^2 + \omega_1^2 x_1^2 + \omega_2^2 x_2^2), \tag{2.1}$$

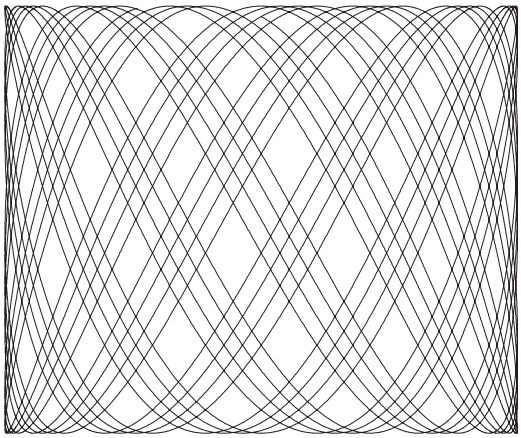
here we have taken units of mass such that $m = 1$, but retained the frequency dependencies. In this treatment we will only consider commensurate frequencies, that is, we will require that ω_1/ω_2 is a rational number.[2] Hamilton's equations are

$$\dot{x}_i = \frac{\partial H}{\partial p_i} \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i} \qquad\qquad i = 1, 2, \tag{2.2}$$

[2]For an elementary introduction to the semiclassics of quantum chaos through the incommensurate harmonic oscillator see [45].



(a) Two color Lissajous figure created by Jonathan Lansey using a Harmonagraph he designed and built.



(b) A Lissajous figure created with a computer. Notice that the orbits alone clearly delineate a rectangular boundary.

Figure 2.1: Lissajous figures

or explicitly,

$$\dot{x}_1 = p_1, \qquad \dot{p}_1 = -\omega_1^2 x_1, \tag{2.3}$$
$$\dot{x}_2 = p_2, \qquad \dot{p}_2 = -\omega_2^2 x_2. \tag{2.4}$$

The general solutions, with initial conditions $(x_{10}, p_{10}, x_{20}, p_{20})$, are

$$x_1(t) = x_{10}\cos(\omega_1 t) + \frac{p_{10}}{\omega_1}\sin(\omega_1 t), \qquad p_1(t) = p_{10}\cos(\omega_1 t) - \omega_1 x_{10}\sin(\omega_1 t), \tag{2.5}$$
$$x_2(t) = x_{20}\cos(\omega_2 t) + \frac{p_{20}}{\omega_2}\sin(\omega_2 t), \qquad p_2(t) = p_{20}\cos(\omega_2 t) - \omega_2 x_{20}\sin(\omega_2 t); \tag{2.6}$$

these are easily implemented on a computer and one can reproduce the art exhibit virtually (see Fig. 2.1(b)). The computer affords the luxury of exploring a wide variety of initial conditions quickly. The analog drawing aesthetically benefits from the damping of the real pendulum, this can be added into your simulations by making the amplitudes $(x_{10}, ..., p_{20})$ time dependent.

These Lissajous figures exhibit the configuration space structure of the oscillator orbits, for our purposes it will also be useful to understand the geometry of the orbits in phase space. Our Hamiltonian has no explicit time dependence and so the total energy,

$$E = E_1 + E_2 = \frac{1}{2}(p_1^2 + \omega_1^2 x_1^2) + \frac{1}{2}(p_2^2 + \omega_1^2 x_2^2) \tag{2.7}$$

is conserved. Clearly the individual oscillator energies are also conserved, and this allows us to take the energy difference as a second independent conserved quantity,

$$K_3 \equiv E_1 - E_2 = \frac{1}{2}(p_1^2 + \omega_1^2 x_1^2) - \frac{1}{2}(p_2^2 + \omega_1^2 x_2^2) \tag{2.8}$$

(the odd subscript 3 is due to the fact that two more oscillator constants will be introduced and will also become more transparent in section 2.3). We can visualize the level sets of these two conserved quantities as follows: specialize to the case $\omega_1 = \omega_2 = 1$ and $E = 1$, then the level set $H = E = 1$,

$$\frac{1}{2}(x_1^2 + p_1^2 + x_2^2 + p_2^2) = 1, \tag{2.9}$$

is a three sphere $S^3 \subset \mathbb{R}^4$. Notation simplifies a bit if we take $z_1 \equiv (x_1 + ip_1)/\sqrt{2}$, $z_2 \equiv (x_2 + ip_2)/\sqrt{2}$, so that the vector (z_1, z_2) is an element of $\mathbb{C}^2 \cong \mathbb{R}^4$. For this vector to lie on the three sphere it must satisfy,

$$|z_1|^2 + |z_2|^2 = 1, \tag{2.10}$$

or in terms of a polar representation $z_1 = r_1 e^{i\xi_1}$, $z_2 = r_2 e^{i\xi_2}$,

$$r_1^2 + r_2^2 = 1. \tag{2.11}$$

This relation is conveniently parametrized in terms of trigonometric functions, let $r_1 = \cos(\theta/2)$ and $r_2 = \sin(\theta/2)$, where the range of θ is $\theta \in [0, \pi]$ in order that r_1 and r_2 are always positive. Of course, these radii are just the partial energies $r_1^2 = E_1, r_2^2 = E_2$ and thus θ parametrizes the energy difference,

$$K_3 = r_1^2 - r_2^2 = \cos^2\left(\frac{\theta}{2}\right) - \sin^2\left(\frac{\theta}{2}\right) = \cos(\theta). \tag{2.12}$$

Now, a point of S^3 is given by

$$(\cos(\frac{\theta}{2})e^{i\xi_1}, \sin(\frac{\theta}{2})e^{i\xi_2}) \qquad \theta \in [0, \pi], \quad \xi_1, \xi_2 \in \mathbb{R}. \tag{2.13}$$

For fixed energy difference K_3, i.e. fixed θ, this is the cartesian product of two circles and so generically describes a torus. Because the oscillator dynamics conserves both E and K_3 the evolution for given initial conditions occurs completely on one of these tori.

A nice way to visualize these tori is by stereographic projection $P : S^3 \subset \mathbb{R}^4 \to \mathbb{R}^3 \cup \{\infty\}$. If we project from the south pole $(0, 0, 0, -1)$ then we have for a point $z = (z_1, z_2) \in S^3$,

$$\begin{aligned} P(z) &= \left(\frac{x_1}{1 + p_2}, \frac{p_1}{1 + p_2}, \frac{x_2}{1 + p_2}\right) \\ &= \left(\frac{\cos(\frac{\theta}{2})\cos(\xi_1)}{1 + \sin(\frac{\theta}{2})\sin(\xi_2)}, \frac{\cos(\frac{\theta}{2})\sin(\xi_1)}{1 + \sin(\frac{\theta}{2})\sin(\xi_2)}, \frac{\sin(\frac{\theta}{2})\cos(\xi_2)}{1 + \sin(\frac{\theta}{2})\sin(\xi_2)}\right) \end{aligned}$$

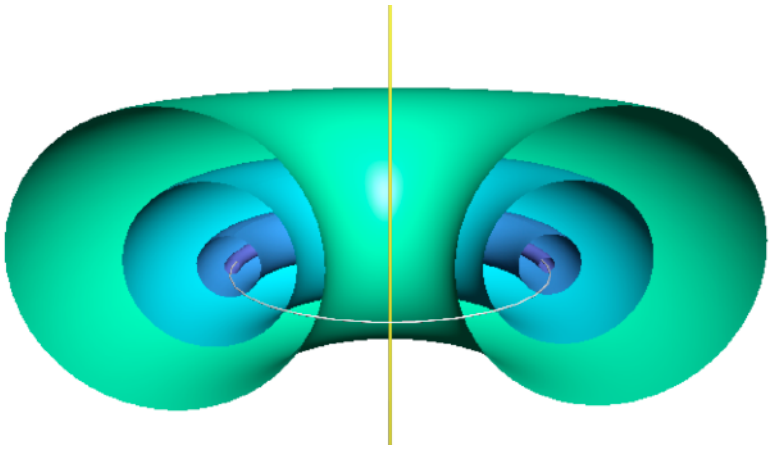


Figure 2.2: A cutaway, side view of several tori in the range $0 < \theta \leq \pi/2$. The white circle represents the degenerate torus as θ goes to 0. The yellow z-axis represents the degenerate torus as θ goes to π.

As we will see below, the value $\theta = \pi/2$ marks a boundary between two qualitatively different sets of tori. As the parameter θ ranges from zero to $\pi/2$ it sweeps out the nested toric leaves of a solid torus $D \times S^1$, here D stands for the topological disk, see Figure 2.2. The parameter value $\theta = 0$ is a circle, viewed in this context as a degenerate torus. The value $\theta = \frac{\pi}{2}$ gives a boundary torus for which $r_1 = r_2 = \frac{\sqrt{2}}{2}$, we'll call this boundary torus T. The parameter values, $\theta \in [\frac{\pi}{2}, \pi]$ also sweep out a solid torus. However these tori degenerate into a circle on S^3 that goes through the south pole and hence is projected onto an entire line, in our coordinates the z-axis. The two degenerate circles are also depicted in Figure 2.2. Because of this projective oddity viewing these tori as nested requires a dual point of view; look at the tori from "inside", see Figure 2.3.

The nesting of the two families of tori can also be visualized by choosing a different point from which to stereographically project, see Figure 2.4. The Tori are also clearly linked with one another from this point of view. This completes our visualization of the level sets (E, K_3).

When the frequency ratio ω_1/ω_2 is rational, $\omega_1/\omega_2 = m_1/m_2$ with m_1 and m_2 integers, the orbits are torus knots, winding one torus cycle m_1 times in the time that the other cycle is traversed m_2 times. The Lissajous figures discussed above are the projection of these orbits onto the configuration space $(\mathbb{R}^2, (x_1, x_2))$. The remarkable fact that these orbits close, that is, that after all this winding the dynamics ends up exactly where it started can be attributed to the fact that there is another quantity conserved by the oscillator dynamics (in our 1:1 resonance case),

$$K_1 = \frac{1}{2}(x_1 x_2 + p_1 p_2). \tag{2.14}$$

This quantity, sometimes called the correlation, is analogous to the Laplace-Runge-

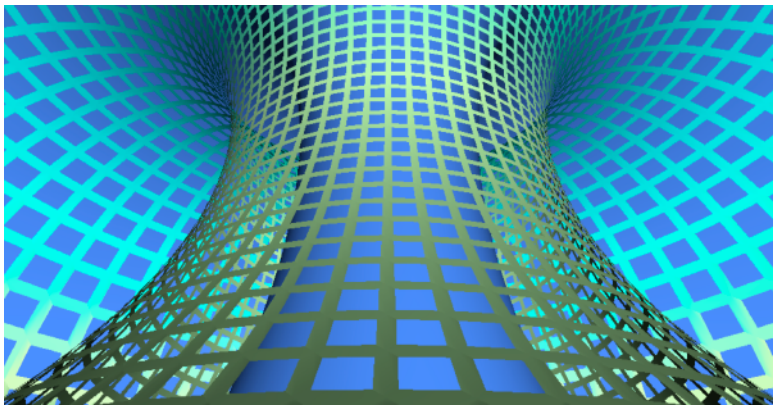


Figure 2.3: The solid torus viewed from "inside." An outer torus has been cut up so that the torus lying beneath it is revealed. Progressing towards deeper tori eventually the degenerate line depicted in yellow in Figure 2.2 is reached.

Lenz vector of the Kepler problem. In particular, the symmetry corresponding to this conserved quantity, a transformation guaranteed by Noether's theorem, is not simply the lift of a configuration space symmetry (see [46]); it is truly a symmetry in the phase space!

The existence of a third, functionally independent conserved quantity has led researchers to call the 2D harmonic oscillator a superintegrable system. Because the integrability language will be used throughout this thesis we briefly review it here.

In the theory of dynamical systems a Hamiltonian system of n degrees of freedom (with a $2n$-dimensional phase space) is said to be integrable if there exist n functionally independent constants of the motion A_i $(i = 1, \ldots, n)$ that are in involution, that is they mutually Poisson commute,

$$\{A_i, A_j\} = 0 \qquad i, j = 1, \ldots, n. \tag{2.15}$$

The celebrated Liouville-Arnold theorem states that the dynamics of an integrable system, for generic cases, occurs completely on an n-dimensional torus which is the level set of these functions. Our 2D oscillator furnishes a nice illustration of this general result. A system is termed superintegrable when it has more than n functionally independent constants of the motion. In this case, not all the conserved quantities can be in involution (otherwise one could take these quantities to be a set of new coordinates $\tilde{q}$ with more than n coordinates). Modern work [47, 48] has demonstrated a close connection between superintegrability and multiseperability, that is the seperability of the Hamilton-Jacobi equations in multiple distinct coordinate systems, but we digress.

Finally we note that the 2D oscillator has a fourth, perhaps more obvious, conserved quantity,

$$K_2 = (x_1 p_2 - x_2 p_1) \tag{2.16}$$

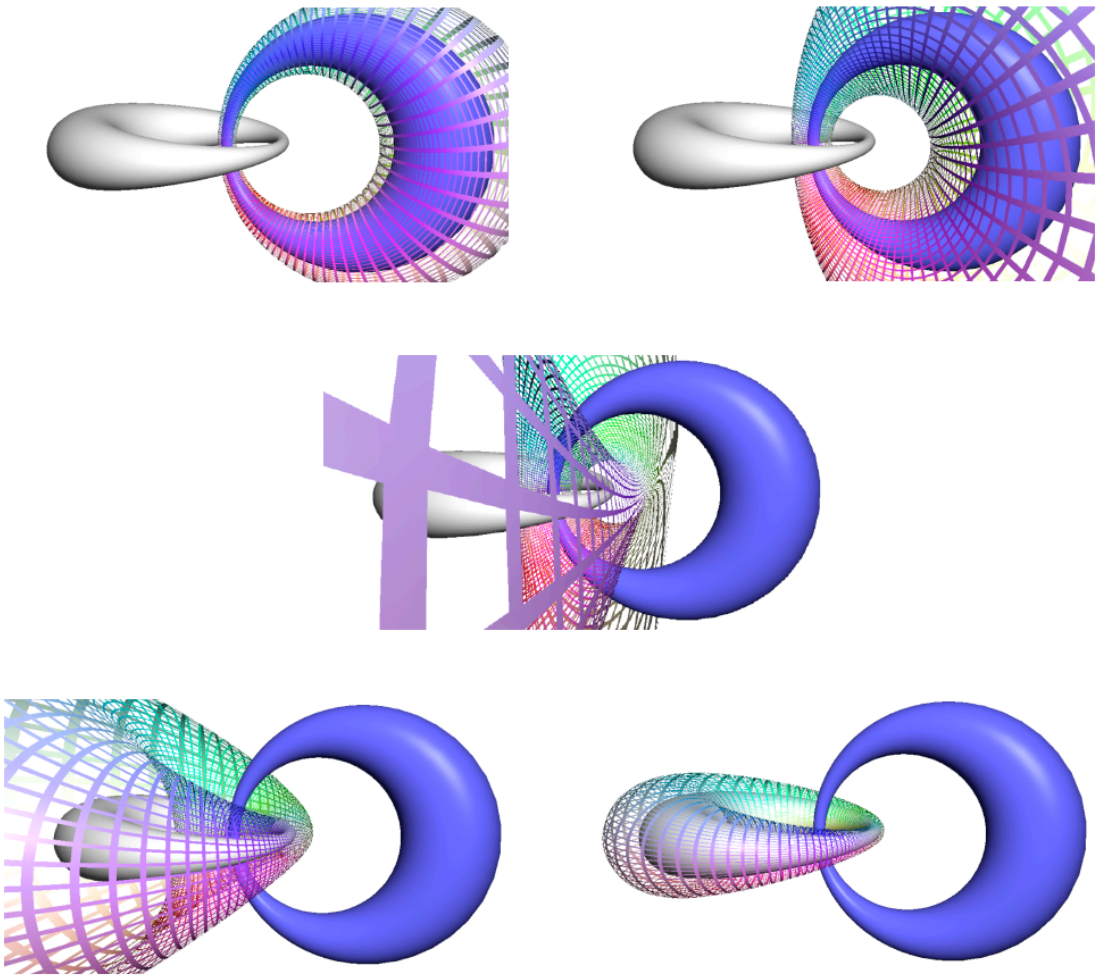


Figure 2.4: Two tori, one from each family are projected after a $\pi/2$ rotation in the p_1p_2 plane. The transition from one family to the other is shown as a cut away torus transitions through various values of θ. In the central picture the cut sheet is the bounding torus T.

(again this is written for the 1:1 resonance). This cannot be independent, else the dynamics would be trivial with the orbit consisting of a single point, and indeed we find,

$$E^2 = K_1^2 + K_2^2 + K_3^2. \tag{2.17}$$

Having outlined the classical geometry of the 2D oscillator, we proceed to semiclassical theory.

2.2 Semiclassics and the 2D oscillator

In this section we explore semiclassics through the 2D harmonic oscillator. The general theory is discussed in several references [41, 42, 38, 43] and will be partially developed below.

2.2.1 WKB theory

The quickest route into WKB theory is to view it as an *ansatz* for solutions to the Schrödinger equation.[3] Assume that the wave function can be written in the form,

$$\psi(\boldsymbol{x}, t) = R(\boldsymbol{x}, t) e^{\frac{i}{\hbar} S(\boldsymbol{x}, t)}, \tag{2.18}$$

with $R(\boldsymbol{x}, t)$ a slowly varying amplitude and the action $S(\boldsymbol{x}, t)$ a more rapidly varying phase.[4] This is not a series expansion for the $\hbar$ dependence of ψ. Indeed, such an expansion would have a physically poor $\hbar \to 0$ limit. The recovery of classical mechanics from quantum theory is more subtle and mathematically more singular. However, it can be thought of as a series for $\ln \psi$,

$$\ln \psi \equiv \frac{i}{\hbar} W(\boldsymbol{x}, t) = \frac{i}{\hbar} \left(W_0(\boldsymbol{x}, t) + W_1(\boldsymbol{x}, t) \frac{\hbar}{i} + W_2(\boldsymbol{x}, t) \left(\frac{\hbar}{i} \right)^2 + \cdots \right), \tag{2.19}$$

with $W_0(\boldsymbol{x}, t) = S(\boldsymbol{x}, t)$ and $W_1(\boldsymbol{x}, t) = \ln (R(\boldsymbol{x}, t))$.

It turns out that this is actually an asymptotic series for the wavefunction and not the more familiar convergent series of first year calculus. Asymptotic series have a qualitatively different behavior, they often arise in circumstances where important partial information is known and pragmatically they give good agreement even at lowest order. However, in general, they do not necessarily improve as higher order

[3] The W, K and B of WKB stand for G. Wentzel, H. A. Kramers and L. Brillouin who all employed this technique at the beginning of the 20th century [49, 50, 51]. Often H. Jeffreys is also added to this list for his early work [52].

[4] Throughout the dissertation we use bold face symbols to denote 3-vectors. We also prefer R (for "radius") over the more conventional choice A (for "amplitude") because A will be pressed into extensive use as we proceed and this choice still retains some mnemonic stickiness through its relation to the polar decomposition.

terms are included. These comments certainly apply to the WKB approximation: the partial input in this case is the form of the WKB wavefunction with its amplitude times phase structure. This approximation often gives quantitative results that are in remarkable agreement with full quantum theory. Also, rather than calculating higher order corrections one is usually led to consider limits where an appropriate quantity with units of action is large with respect to $\hbar$, as this quantity grows the approximation becomes more and more accurate. Often this sort of limit can be interpreted as giving rise to a classical theory, although not always, and it is really in this sense that researchers consider the oft quoted $\hbar \to 0$ limit. The relation to asymptotic series also explains why physicists frequently refer to this area as the study of asymptotics. If greater accuracy is desired even at small actions then hyperasymptotics can be pursued, although this can be intractable analytically, see the excellent review by Boyd [53]. Other asymptotic series familiar in physics are the Sterling approximation and Feynman diagram expansions (there are many more examples in [53]).

Putting the *ansatz* (2.18) into the Schrödinger equation one finds, at lowest orders, that $S(\boldsymbol{x}, t)$ and $R(\boldsymbol{x}, t)$ should satisfy the two equations:

$$\frac{1}{2m}(\boldsymbol{\nabla} S)^2 + V(\boldsymbol{x}, t) + \frac{\partial S}{\partial t} = 0 \tag{2.20}$$

$$\frac{1}{2m} R \boldsymbol{\nabla}^2 S + \frac{1}{m} \boldsymbol{\nabla} S \cdot \boldsymbol{\nabla} R + \frac{\partial R}{\partial t} = 0, \tag{2.21}$$

here $V(\boldsymbol{x}, t)$ is the potential for the system of interest. The first of these two equations is the Hamilton-Jacobi equations for $S(\boldsymbol{x}, t)$,

$$H(\boldsymbol{x}, \boldsymbol{\nabla} S, t) + \frac{\partial S}{\partial t} = 0, \tag{2.22}$$

where $H(\boldsymbol{x}, \boldsymbol{p}, t)$ is the classical Hamiltonian. There will be more to say on the implied identification $\boldsymbol{p} = \boldsymbol{\nabla} S$ briefly. The second of these equations depends on both R and S. Practically then, one proceeds by first finding S, i.e. by solving (2.22), and then solving for the amplitude, R. The equation (2.21), which we will call the amplitude transport equation, can also be cast in a more familiar form. Let $\boldsymbol{v}(\boldsymbol{x}, t) \equiv 1/m \boldsymbol{\nabla} S(\boldsymbol{x}, t)$ and $\rho(\boldsymbol{x}, t) \equiv R^2(\boldsymbol{x}, t)$ then, multiplying (2.21) through by R and rearranging yields the equivalent,

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v}) = 0. \tag{2.23}$$

This is a continuity equation for the probability density $\rho = R^2 = |\psi|^2$ and so represents the conservation of total probability.

Beginning with an *ansatz*, in our case $\psi(\boldsymbol{x}, t) = R(\boldsymbol{x}, t) \exp\{\frac{i}{\hbar} S(\boldsymbol{x}, t)\}$, is the mathematical physicist's version of a rabbit from the hat. However, there is a nice physical justification that can be given and it also motivates the identification $\boldsymbol{p} = \boldsymbol{\nabla} S$

encountered above. The physical assumption is that we are interested in a wave phenomenon. In particular, one for which the amplitude is slowly varying in comparison to the wavelength. In these circumstances we can define a local wavelength, λ, that is just the distance between successive wave fronts (for simplicity consider 1D),

$$S(x+\lambda,t) = S(x,t) + 2\pi\hbar. \tag{2.24}$$

Taylor expanding about x on the left and simplifying yields,

$$\frac{\partial S}{\partial x}\lambda = 2\pi\hbar \qquad \text{or} \qquad \frac{\partial S}{\partial x} = \frac{2\pi\hbar}{\lambda}. \tag{2.25}$$

The λ is only well defined locally and so we view it as a function $\lambda(\boldsymbol{x},t)$ depending on location and time. Finally, interpreting the right hand side in terms of de Broglie's relation $p = h/\lambda$ (and generalizing to 3D) we have,

$$\boldsymbol{\nabla} S = \boldsymbol{p}(\boldsymbol{x},t), \tag{2.26}$$

with $p(\boldsymbol{x},t) = h/\lambda(\boldsymbol{x},t)$. The wavefunction behaves locally like a wave or swarm of classical particles each one obeying the de Broglie hypothesis.[5] A similar analysis can be performed on the t variable,

$$S(\boldsymbol{x},t+T) = S(\boldsymbol{x},t) - 2\pi\hbar, \tag{2.27}$$

here T is the period and the minus sign is a convention (think of a plane wave $\exp(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)$). Expanding yields,

$$\frac{\partial S}{\partial t}T = -2\pi\hbar \qquad \text{or} \qquad \frac{\partial S}{\partial t} = -\omega\hbar. \tag{2.28}$$

This time we introduce Einstein's relation $E = \hbar\omega$ to obtain,

$$-\frac{\partial S(\boldsymbol{x},t)}{\partial t} = E(\boldsymbol{x},t), \tag{2.29}$$

and again $E(\boldsymbol{x},t)$ is interpreted as a local energy. Then the Hamilton-Jacobi equation,

$$-\frac{\partial S}{\partial t} = H(\boldsymbol{x},\boldsymbol{\nabla}S,t), \tag{2.30}$$

tells us that $\boldsymbol{p} = \boldsymbol{\nabla}S$ and $E(\boldsymbol{x},t) = -\partial S/\partial t$ are not independent, instead they are related by $H(\boldsymbol{x},\boldsymbol{p}(\boldsymbol{x},t),t) = E(\boldsymbol{x},t)$.

The amplitude transport equation can also be understood as the evolution of the density of this swarm of particles and in particular as imposing the conservation of the total number of particles, see [38]. This concludes our brief introduction to the physical intuition behind the WKB approximation.

[5] If this picture reminds you of Huygens' principle in optics then you are in good company. Schrödinger strongly emphasized this analogy in his derivation of the wave equation [54, 55] (the second reference contains an english translations of all four of Schrödinger's communications on his wave equation).

2.2.2 Lagrangian manifolds

Standard treatments focus on WKB theory for systems of one degree of freedom. While one degree of freedom applications will play an important role in what follows we will also be treating systems that are essentially higher dimensional. The development of WKB theory to this more general context is also of interest in itself. We will find that there is an exquisite geometrical core to the theory; this core is the subject of Lagrangian manifolds. It is hardly an exaggeration to say that semiclassics is the study of Lagrangian manifolds. The geometrical framework presented here will not only ease the application of the technique but also enrich its physical content.

There is a remarkable consequence of the demonstration from the last section that the phase S can be interpreted as a solution of the Hamilton-Jacobi equation: such a solution picks out a submanifold of the phase space, and this manifold is always Lagrangian. The manifold is given by the graph, that is $(q, p(q))$, of the momentum field $\boldsymbol{p} = \boldsymbol{\nabla} S$ associated to S.[6] Observe that momentum fields arising from such an S are not general, these momentum fields are always curl free,

$$\boldsymbol{\nabla} \times \boldsymbol{p} = \boldsymbol{\nabla} \times \boldsymbol{\nabla} S = 0. \tag{2.31}$$

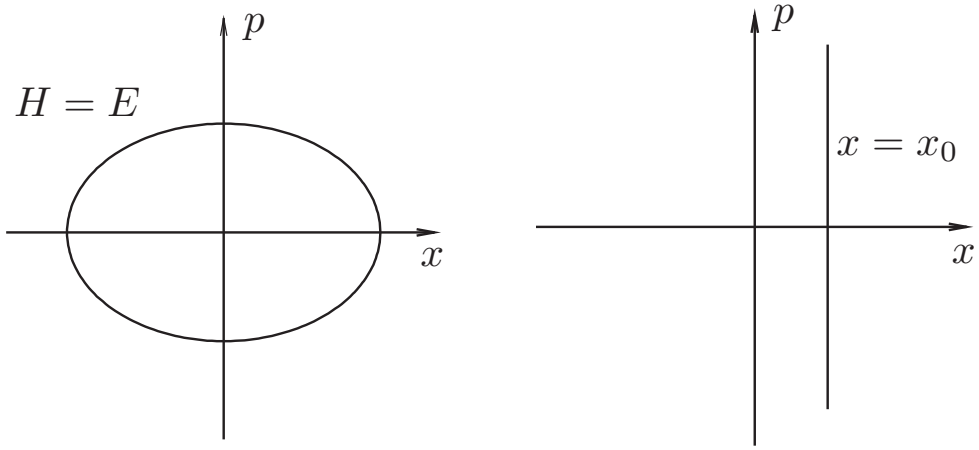
This observation is the simplest version of the more abstract condition defining a Lagrangian manifold that will be developed below. Lagrangian manifolds are n-dimensional submanifolds of the $2n$-dimensional phase space and can be thought of as semiclassical states. By a semiclassical state we mean a classical object that as closely mimics a quantum state as possible. Every aspect of the semiclassical wavefunction and the associated eigenvalues is built around these manifolds.

Before turning to the development of the notion of a Lagrangian manifold let us describe some examples. In the Hamiltonian framework, Lagrangian manifolds can be specified as the level set of n Poisson commuting phase space functions (observables) $A_i(\boldsymbol{x}, \boldsymbol{p})$ $(i = 1, \ldots, n)$. Because the phase space functions A_i are commuting they act like a set of generalized coordinates or momenta (recall that there is a canonical transformation that switches the two). One way to think of Lagrangian manifolds, then, is as very general (usually curvilinear) coordinate spaces for the wave function. The tori discussed in section 2.1 are Lagrangian submanifolds of the 2D oscillator phase space (see Fig. 2.2). They are specified by the level set of the two functions $A_1 = H$ and $A_2 = K_3$, which satisfy $\{H, K_3\} = 0$.

The simplicity of the 2D oscillator is partially derived from the fact that the motion is completely separable into two oscillators. Another example of a Lagrangian

[6] When there is little risk of confusion we abbreviate dependencies on the vector of coordinates and momenta $(q_1, \ldots, q_n)$ and $(p_1, \ldots, p_n)$ to q and p. Also, at this point we transition to using q for configuration space coordinates. This is intended to mark the fact that these may be generalized coordinates. We will use x henceforth to emphasize consideration of the configuration space $\mathbb{R}^n$ and q will be used when the configuration space is more general but also, when convenient, in the Euclidean case.

manifold is the constant energy ellipse in the phase plane of one of these oscillators, see Fig. 2.5(a). An interesting example is pictured in Fig. 2.5(b). This Lagrangian



(a) Lagrangian manifold of constant energy.

(b) Vertical Lagrangian manifold with constant x.

Figure 2.5: Examples of Lagrangian manifolds.

manifold is specified by a fixed value of the x coordinate, $x = x_0$. Notice that in this case, there is no way in which a momentum field can be associated to the Lagrangian manifold because the manifold is not describable by a function $p(x)$. Lagrangian manifolds that are vertical over the configuration space are one important example that illustrates why we will need to generalize the condition (2.31).

There is a very simple physical intuition behind the interpretation of Lagrangian manifolds as semiclassical states. In full quantum theory the Heisenberg uncertainty principle says that there is a limit to the accuracy with which one can simultaneously know the value of an observable $\hat{A}$ and the value of the conjugate variable, call it $\hat{\alpha}$. The most famous example, of course, being $\Delta\hat{x}\Delta\hat{p} \geq \hbar/2$. In quantum theory we are often interested in eigenstates of an operator, say $\hat{A}$, and these states have a sharp value, say a, of this observable, the eigenvalue. This means that they should be very spread in the conjugate variable $\hat{\alpha}$. This is precisely how Lagrangian manifolds behave. Take the two one-dimensional examples introduced above. In Figure 2.5(a) the oscillator has a definite energy but if all we're given is the geometrical manifold we know nothing about the parameter along the manifold, that is, the time in this example. This is even simpler to see in the second example (Fig. 2.5(b)) which has a sharp value of position $x = x_0$ but is completely spread in momentum.

Lagrangian manifolds are properly understood in the geometrical formulation of mechanics, which we briefly describe now, see [56, 57, 58]. A physical system is determined firstly by the specification of its kinematics, the arena of possibilities

embodied by the phase space; and secondly by its dynamics embodied in the choice of a Hamiltonian, $H(q,p)$. The evolution is determined by a combination of both of these structures. On the kinematical side, once an n-dimensional configuration space Q has been specified, with an n-tuple of coordinates q, then there is a phase space $\Phi \equiv T^*Q$ that can always be constructed having Q as its configuration space. This is called the cotangent bundle, a fancy name for phase space, and is coordinatized by the $2n$-tuple (q,p). The important point here is that the phase space comes endowed with a natural structure, the symplectic two form ω, that can be used to convert phase space functions into vector fields (for details on the construction of the cotangent bundle and the symplectic form see Arnold's wonderful book [56]). In coordinates the symplectic form is simply

$$\omega = \sum_{i=1}^{n} dp_i \wedge dq^i, \tag{2.32}$$

and Darboux's theorem guarantees that, at least locally, it can always be written in this form. The symplectic form is a field defined at each point $(q,p) \in T^*Q$, this won't be belabored except where necessary. The symplectic form can be used to associate to $H(q,p)$ the Hamiltonian vector field X_H. The Hamiltonian vector field is defined by the relation,

$$\omega(\cdot, X_H) = dH(\cdot). \tag{2.33}$$

Informally, the symplectic form, a two form, eats a vector and returns a one form; when fed the Hamiltonian vector field it returns the differential of the Hamiltonian. In the standard coordinate basis for vector fields $(\partial/\partial q, \partial/\partial p)$, X_H has components (X_{Hq}, X_{Hp}), that is,

$$X_H = X_{Hq}\frac{\partial}{\partial q} + X_{Hp}\frac{\partial}{\partial p}. \tag{2.34}$$

These components are determined by (2.33),

$$\omega(\cdot, X_H) = -X_{Hp}dq + X_{Hq}dp = dH(q,p) = \frac{\partial H}{\partial q}dq + \frac{\partial H}{\partial p}dp, \tag{2.35}$$

and thus,

$$X_H = \frac{\partial H}{\partial p}\frac{\partial}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial}{\partial p}. \tag{2.36}$$

With the Hamiltonian vector field in hand the construction of the system's evolution is the same as the construction of the electric field lines from a given electric field. Geometrically, this amounts to locally following the vector field and constructing its integral curves. Analytically this is achieved by equating the Hamiltonian vector field with the tangent vector of the time evolution at each point,

$$\dot{q} = X_H q = \frac{\partial H}{\partial p} \qquad \text{and} \qquad \dot{p} = X_H p = -\frac{\partial H}{\partial q}, \tag{2.37}$$

yielding Hamilton's equation! At times, too little emphasis is placed on the fact that one needs both ω and $H(q,p)$ in order to construct a system's evolution.

Lagrangian manifolds are defined in terms of the symplectic structure. A few technical conditions are necessary to ensure that the symplectic form is well defined. As a two form, ω is antisymmetric and we will assume that it is closed, $d\omega = 0$, and non-degenerate, that is, if

$$\omega(X, Y) = 0, \tag{2.38}$$

for a fixed vector X and all vectors Y then the vector X is zero. At last, we are ready to provide the full definition of a Lagrangian manifold: a Lagrangian manifold L is an n-dimensional submanifold of the phase space such that for any two vectors X_1 and X_2 tangent to L we have,

$$\omega(X_1, X_2) = 0. \tag{2.39}$$

In coordinates this becomes, for the case of a 2-dimensional phase space,

$$X_{1p}X_{2q} - X_{1q}X_{2p} = 0, \tag{2.40}$$

where $X_1, X_2 \in T_{(q,p)}L \subset T_{(q,p)}\Phi$ and both vectors have been expressed in the coordinate basis $X_1 = X_{1q}\partial/\partial q + X_{1p}\partial/\partial p, X_2 = X_{2q}\partial/\partial q + X_{2p}\partial/\partial p$. And in the general case,

$$X_{1p_i}X_{2q^i} - X_{1q^i}X_{2p_i} \equiv X_{1p} \cdot X_{2q} - X_{1q} \cdot X_{2p} = 0, \tag{2.41}$$

with $X_1 = X_{1q^i}\partial/\partial q^i + X_{1p_i}\partial/\partial p_i$ and similarly for X_2. This is a local condition that precisely captures the idea that the manifold L looks like pure configuration space for some choice q of generalized coordinates with no momenta mixed in. As mentioned above, there is a canonical transformation that switches q's and p's and so L may be viewed as pure momentum space too. The condition (2.39) is equivalent to the involution of the functions whose level set is the manifold and so we recover the definition in terms of commuting observables near the beginning of this section.

Having developed an analytic characterization in terms of the symplectic form we return to the examples introduced above. Observe that for phase spaces Φ of one degree of freedom (two dimensions) the condition (2.40) is trivial. Because for any one dimensional submanifold of Φ the tangent space at a point is one dimensional, with basis E say, and hence all tangent vectors are linearly dependent, so $\omega(X_1, X_2) = c\omega(E, E) = -c\omega(E, E) = 0$. In higher dimensions the condition becomes nontrivial. Two immediate generalizations of the example 2.5(b) are the n-dimensional surfaces of constant q or constant p. These are Lagrangian because $\omega(X_1, X_2)$ for vectors X_1 and X_2 that have only $\partial/\partial p$ or $\partial/\partial q$ components vanishes. These examples are important because they are the analogs of the position and momentum bases $|x\rangle$ and $|p\rangle$ in quantum theory. Wavefunctions, e.g. $\psi(x) = \langle x|\psi\rangle$, are generally evaluated in one of these bases and this will continue to be the case in semiclassics where these special Lagrangian manifolds will play an important role in evaluating EBK wavefunctions in position or momentum bases.

This section began with the observation that momentum fields arising from the Hamilton-Jacobi equation were curl free but thus far no connection between this fact and the notion of Lagrangian manifolds has been drawn. Consider a surface constructed from the solution S to the Hamilton-Jacobi equation given by the graph $(q, p(q))$, where as always $p_i = \partial S/\partial q^i$. On this surface the differentials dq^i and dp_i are not independent, in particular $dp_i = (\partial p_i/\partial q^j)dq^j$. This implies that the components of a vector tangent to this graph, $X = X_{q^i}\partial/\partial q^i + X_{p_i}\partial/\partial p_i$, are related by,

$$dp_i(X) = X_{p_i} = \frac{\partial p_i}{\partial q^j} X_{q^j} = \frac{\partial p_i}{\partial q^j} dq^j(X). \tag{2.42}$$

On the other hand the fact that p is curl free can be written in coordinates as,

$$\frac{\partial p_i}{\partial q^j} - \frac{\partial p_j}{\partial q^i} = 0. \tag{2.43}$$

A short calculation using (2.42) and (2.43) leads to the conclusion that momentum fields arising from solutions to the Hamilton-Jacobi equation have graphs $(q, p(q))$ that are Lagrangian!

The relationship between Lagrangian manifolds and the actions S can be further elucidated now. As already noted, not all Lagrangian manifolds can arise from curl-free momentum fields. The trouble is that $p(q)$ may not be single valued or even a well-defined function, both examples of one dimensional Lagrangian manifolds discussed above illustrate this point (Fig. 2.5). The vertical Lagrangian manifold is a particularly singular case for which there are infinitely many p's above the point x_0. These singular cases are the exception and for most regions of configuration space it is possible to label the different branches of the inverse projection that takes a point of configuration space to the Lagrangian manifold with a branch index $(k = 1, \ldots, l)$, where l is the total number of branches. These branches label independent solutions of the Hamilton-Jacobi equation S_k. In general then, most Lagrangian manifolds arise as the gluing together of l momentum fields determined by the l branches S_k. In the example illustrated in Figure 2.5(a) the two branches of the momentum come about in a familiar manner as the two branches of the square root function,

$$p_1 = \pm\sqrt{2E - \omega_1^2 x_1^2}. \tag{2.44}$$

A mildly restricted converse to the result discussed above that curl-free momentum fields give rise to Lagrangian manifolds can be given, see [38] for details. Focus on a single branch k of a Lagrangian manifold (one that projects in a non-singular fashion onto the configuration space), choose an arbitrary initial point in this surface and calculate,

$$S_k(q) = \int^q p(q)dq, \tag{2.45}$$

along a path confined to the Lagrangian manifold. Modulo topological considerations, which can give rise to additional multivaluedness of the action, this integral is path independent and yields a curl-free momentum field $p_i = \partial S/\partial q^i$. Topological considerations will be relevant in applications because, as has already been pointed out above, the Liouville-Arnold theorem guarantees that in generic cases integrable systems are associated to tori. Non-trivial topologies will be treated on a case by case basis.

The transitions between the various branches that glue together to form a Lagrangian manifold are an important facet of semiclassical theory. Consider the vertical projection $\pi : L \to Q$ that takes a point $(q, p) \in L$ to the configuration space point lying under it, $\pi(q, p) = q$. As discussed above, for points q lying under L, the inverse of this projection is generally multivalued and its branches are the various sheets of the Lagrangian manifold. When two branches meet the character of this projection changes. In particular, for points not at such a join, the tangent map $T\pi : TL \to TQ$, or in physicists language the Jacobian of π, maps the tangent plane of L in a regular manner (smoothly) to the tangent plane of Q. However, at a join the tangent plane becomes vertical and so $T\pi$ has a non-trivial kernel, in other words the Jacobian matrix of π is not full rank. The coordinate calculations underlying this geometrical picture will be discussed in greater detail for the 2D oscillator below, section 2.2.3. The subset of points of the Lagrangian manifold at which the tangent planes are vertical will be called the critical set and its image under π in Q is defined as the caustic set (see Fig. 2.6). The name critical set for the points on the Lagrangian manifold originates from a slightly broader perspective. The Lagrangian manifold is equally well described in another representation, for example the momentum representation where it is locally described by a position field $q(p)$. The critical set on the Lagrangian manifold for the q projection π is precisely the critical set of this function $q'(p) = 0$. A readable treatment of the representation dependence of the caustic set that also sets up applications of caustics in general relativity is Ehlers and Newman's review [59].

Again these generalities are simple and familiar for the oscillator (Figs. 2.7(a) and 2.7(b)). The caustic points are the points at which the wavefunction transitions from its standard oscillatory form to the damped evanescent wave. On the classical side these are the classical turning points of the motion. An oscillator with fixed energy E cannot go beyond them without violating conservation of energy.

Geometrically it is clear that the caustic set is representation dependent. The projection $\tilde{\pi} : \Phi \to P$, here P is the subspace spanned by the momentum coordinates p_i, will see different branches and their joins. Thus the caustic set really reveals the relationship between the basis used and the state under consideration.

This brief introduction to Lagrangian manifolds provides the necessary background for understanding the structure of semiclassical wavefunctions that we turn to now.

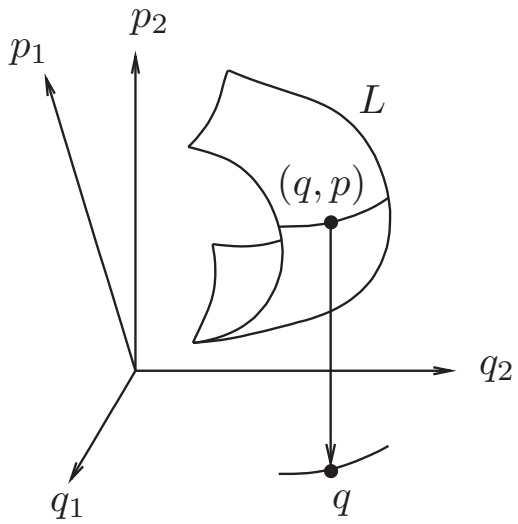
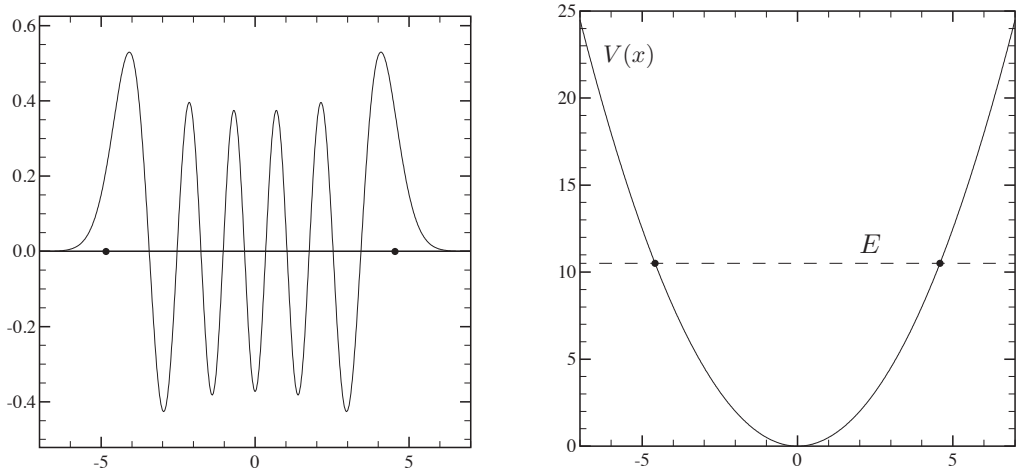


Figure 2.6: The caustic set of L is depicted in the q_1-q_2 plane. This set arises from the projection of the critical set on the Lagrangian manifold where the tangent space to the manifold is vertical over q-space in one or more directions. Figure courtesy of R. G. Littlejohn.



(a) Normalized harmonic oscillator wavefunction with quantum number $n = 10$. The two dots on the x-axis mark the caustic points of the Lagrangian manifold depicted in Fig. 2.5(a).

(b) The harmonic oscillator potential (solid line) and the level set of the energy (dashed line) corresponding to the wavefunction depicted at left, 2.7(a). The classical turning points at this energy are marked by dots.

Figure 2.7: The caustics of the harmonic oscillator (left) and the corresponding classical turning points (right).

2.2.3 Semiclassics of the 2D oscillator

As was alluded to in the introduction to this chapter, we will not begin with the general theory and specialize to examples later. Instead, in this section we will explore the example of the 2D oscillator in detail. Throughout this section we will repeatedly rely on the separability of the 2D oscillator into two independent oscillators to simplify calculations and visualizations; no further explicit mention of this property will be made. By the end of the section the general results will be well motivated and easy to describe.

The promethean problem of semiclassics is the establishment of a correspondence between a quantum system and a classical counterpart. Quantization (and dequantization also) is part science and part art. If this weren't true, gravity would have been quantized some time ago. When the Hilbert space of interest is $L^2(\mathbb{R}^n)$ there is a well defined prescription for associating a classical phase space function to quantum operators, this is the Weyl transform. The classical "symbol" of the operator $\hat{A}$ is generally an even power series in $\hbar$ and the $\hbar$ independent term is called the principal symbol, it will be denoted by $A(q,p)$ with no hat. Our focus will be on the lowest order quantum effects and the terms of order $\hbar^2$ will be systematically ignored. The result is that the operator-symbol correspondence is straightforward in everything that follows: express the operator in terms of the fundamental operators $\hat{q}$ and $\hat{p}$ and then remove the hats.

The Hilbert space of the 2D oscillator is $L^2(\mathbb{R}^2)$ with standard wavefunctions $\psi(x_1, x_2)$, the eigenstates of the total energy,

$$\hat{H} = \frac{1}{2}(\hat{p}_1^2 + \omega_1 \hat{x}_1^2 + \hat{p}_2^2 + \omega_2 \hat{x}_2^2), \tag{2.46}$$

and the energy difference,

$$\hat{K}_3 = \frac{1}{2}(\hat{p}_1^2 + \omega_1 \hat{x}_1^2 - \hat{p}_2^2 - \omega_2 \hat{x}_2^2). \tag{2.47}$$

The eigenvalues of these operators will be labelled with quantum numbers n and m. States can also be labelled by the eigenvalues n_1 and n_2 of the partial energies $\hat{H}_1 = 1/2(\hat{p}_1^2 + \omega_1 \hat{x}_1^2)$ and $\hat{H}_2 = 1/2(\hat{p}_2^2 + \omega_2 \hat{x}_2^2)$.

Except for a small part of Chapter 4 we will be concerned exclusively with the semiclassics of systems whose classical counterparts are integrable. The importance of this restriction was first appreciated by Einstein [60] and later developed by Brillouin and Keller [61, 62] and for this reason it is usually called EBK quantization. In particular this has an impact on the semiclassical quantization conditions, which generalize the Bohr-Sommerfeld rule of the old quantum theory. We will discuss this in detail briefly.

The generalization of the WKB wavefunctions of section 2.2.1 to multidimensional integrable systems has a remarkably similar form. It consists of a sum of terms each

one just like the WKB *ansatz*, for the 2D oscillator we have,

$$\psi(x_1, x_2) = \frac{i}{\sqrt{V_A}} \sum_{k=1}^{4} R_k \exp\left\{ \frac{i}{\hbar} \left[S_k(x, a) - \mu_k \frac{\pi}{2} \right] \right\}. \tag{2.48}$$

Notation is explained as we proceed. The two most novel features here are the phase adjustments μ_k which are called Maslov indices and are required to be positive or negative integers and the V_A which is a volume factor whose origin is explained momentarily. The sum runs over the distinct intersections of two Lagrangian manifolds: the constant x manifold with (x_1, x_2) fixed but arbitrary and the Lagrangian manifold given by constant H and K_3, one of the tori of section 2.1. These are each 2-dimensional submanifolds of $\Phi = \mathbb{R}^4$ and so when they intersect they generically do so in a discrete set of points. (Recall that the dimension of the generic, more precisely transverse, intersection I of two submanifolds M_1 and M_2 is determined by the codimension sum rule, $\operatorname{codim} I = \operatorname{codim} M_1 + \operatorname{codim} M_2$. Thus for two Lagrangian manifolds we have $\dim \Phi - \dim I = 2 \dim \Phi - \dim M_1 - \dim M_2$ or $\dim I = \dim M_1 + \dim M_2 - \dim \Phi = 0$.) There are four such intersections and they can be labelled by four distinct signed combinations of momenta: consider the first oscillator, intersecting $x_1 = x_{10}$ and the constant energy surface $H_1 = E_1$ gives the two

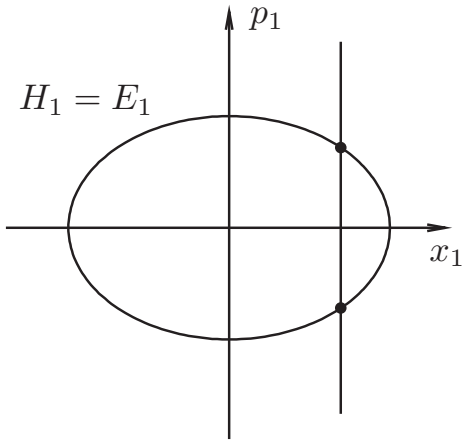


Figure 2.8: The two intersection points of the Lagrangian manifolds $x_1 = x_{10}$ and $H_1 = E_1$ are at the momentum values $\pm p_{10}$.

points $\pm p_{10}$ (Fig. 2.8). The same is true of the second oscillator and so the four intersections above (x_{10}, x_{20}) are $\{(p_{10}, p_{20}), (-p_{10}, p_{20}), (p_{10}, -p_{20}), (-p_{10}, -p_{20})\}$. This appears in a very geometrical fashion on the Lissajous figure. Given a Lissajous figure (Fig. 2.9) there are several points at which the orbit crosses itself. The dynamics is consistent with all four vectors that are tangent to the orbit at this point. This can be

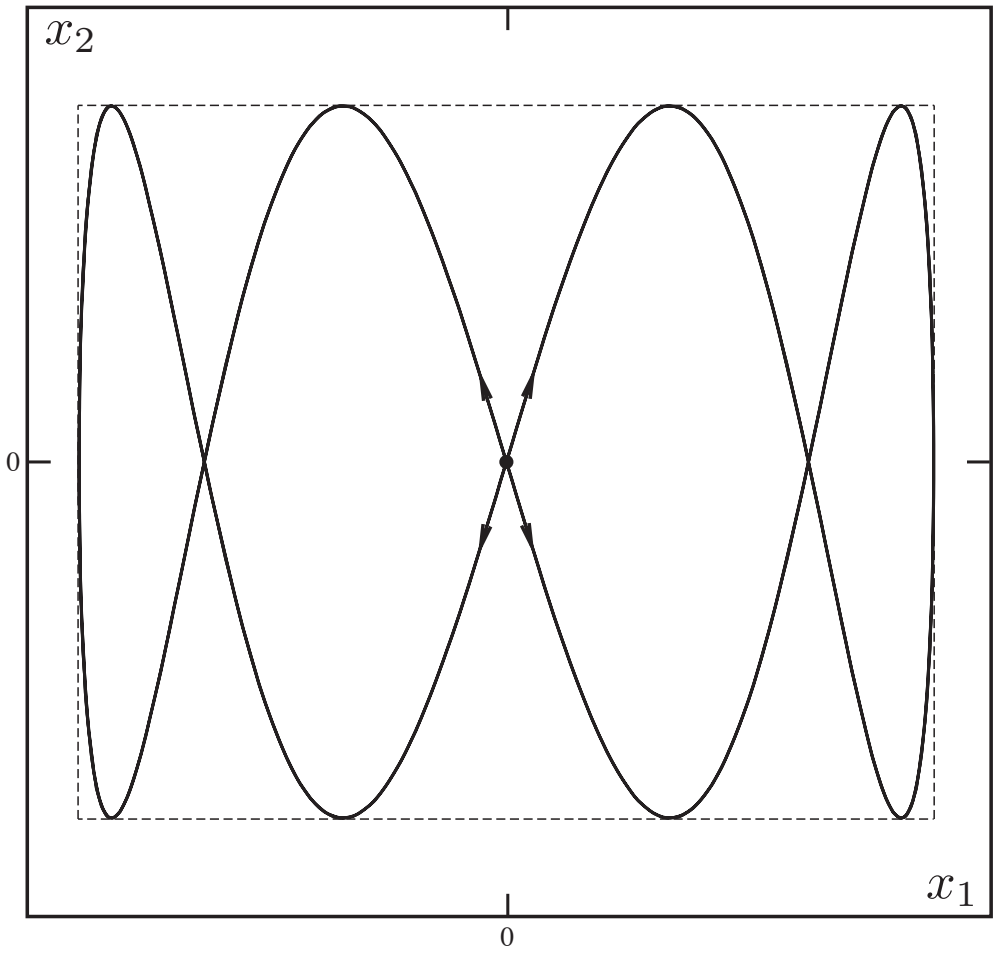


Figure 2.9: Lissajous figure illustrating the projection of the four branches of the 2D oscillator Lagrangian manifold $H = E$ and $K_3 = k_3$ onto configuration space. The dashed lines are caustics, on the vertical dashed lines $p_1 = 0$ and on the horizontal ones $p_2 = 0$. At each of the corners of the dashed rectangle all four branches join into a single point sitting above the configuration space.

demonstrated as follows: trace the orbit along the direction specified by one of these tangent vectors, upon returning to the intersection point the evolution will be in one of the other directions. However, the oscillator is time reversal invariant and so upon reversing time (or more physically momenta of the initial conditions) the other two vectors are seen to also be consistent with the orbit. These then are the four branches of the 2D oscillator Lagrangian manifold as it sits above the (x_1, x_2) configuration space. The coordinate calculations of section 2.1 have already demonstrated that these branches glue together to form a torus. However, an interesting perspective on this can be gained by examining the amplitudes of the wavefunction (2.48).

An under appreciated but computationally invaluable formula for the amplitudes R_k was found in [63]. The amplitude at the intersection of two Lagrangian manifolds specified by two complete sets of commuting observables $A = \{A_1, \ldots, A_n\}$ and $B = \{x_1, \ldots, x_n\}$ is given by $R_k = |D_k|^{-1/2}$ with,

$$D_k \equiv \det \{x_i, A_j\}, \tag{2.49}$$

we will call D the amplitude determinant. The poisson brackets are evaluated at the intersection of the A-manifold with the constant x plane on the kth branch of the inverse projection from x-space to the A Lagrangian manifold. This is equivalent to the usual expression for the amplitude in terms of a determinant of the Hessian of the phase S_k,

$$D_k^{-1} = \det \frac{\partial^2 S_k(x,a)}{\partial x_i \partial a_j}, \tag{2.50}$$

but in many circumstances can be vastly simpler to calculate. Here the action $S_k(x,a)$ is the projection of the kth branch of the action S on the Lagrangian manifold determined by $\{A_i\}_{i=1,\ldots,n} = \{a_i\}_{i=1,\ldots,n} = a$ to the configuration space. For the 2D oscillator the set of A observables is $A = \{A_1, A_2\} = \{H, K_3\}$ and it is useful for calculating Poisson brackets to note that $H = H_1 + H_2$ and $K_3 = H_1 - H_2$.[7] Then the amplitude determinant is,

$$D = \det \begin{pmatrix} \{x_1, H\} & \{x_1, K_3\} \\ \{x_2, H\} & \{x_2, K_3\} \end{pmatrix} = \det \begin{pmatrix} p_1 & p_1 \\ p_2 & -p_2 \end{pmatrix} = -2p_1 p_2, \tag{2.51}$$

and so,

$$R_k = \frac{1}{\sqrt{2|p_1 p_2|}}, \tag{2.52}$$

evaluated on the appropriate branch. The functional form of this amplitude is no surprise, it's just the product of the two individual oscillator amplitudes $1/\sqrt{p_1}$ and $1/\sqrt{p_2}$. The factor of $1/\sqrt{2}$ here and the new factor, $1/\sqrt{V_A}$ multiplying the right hand side of (2.48) combine to give the proper normalization to the wavefunction. The latter is the inverse square root of the volume of the torus defined by the A

[7] Note that we use curly braces both for Poisson brackets and for sets.

observables, measured with respect to the density $d\alpha_1 \wedge \cdots \wedge d\alpha_n$ in the conjugate variables α.

This point can require some care. In our example, let the variable conjugate to H be t and that conjugate to K_3 be ϕ then,

$$V_A = \int \int_L dt \wedge d\phi, \tag{2.53}$$

where L is the Lagrangian manifold $H = E$ and $K_3 = k_3$. Because the periods and region of integration are more obvious in the variables t_1 and t_2 conjugate to H_1 and H_2 it is convenient to switch to these variables. The generating function of type 2 (old positions and new momenta) for this transformation is given by $F(t_1, t_2, E, K_3) = 1/2[t_1(E + K_3) + t_2(E - K_3)]$. It is an easy check that $\partial F/\partial t_1 = E_1$ and $\partial F/\partial t_2 = E_2$ while,

$$t = \frac{\partial F}{\partial E} = \frac{1}{2}(t_1 + t_2) \qquad \phi = \frac{\partial F}{\partial K_3} = \frac{1}{2}(t_1 - t_2). \tag{2.54}$$

This leads to a Jacobian factor for the change of coordinates $(t, \phi) \mapsto (t_1, t_2)$ of $1/2$. So,

$$V_A = \int \int_L dt \wedge d\phi = \frac{1}{2} \int \int_L dt_1 \wedge dt_2 = \frac{1}{2} \int_0^{2\pi/\omega_1} \int_0^{2\pi/\omega_2} dt_1 dt_2 = \frac{2\pi^2}{\omega_1 \omega_2}. \tag{2.55}$$

An interesting observation about the amplitudes R_k is that they can also be used to find the caustic; these amplitude factors diverge precisely on the caustic. This well known singularity of WKB theory forces a careful analysis of behavior near the caustics and, in differential equations language, can be smoothed out by developing connection formulae [64].[8] The appearance of this singularity at the cuastic has a nice geometrical interpretation that is facilitated by the formula for the amplitude determinant (2.49). This determinant is zero when the rank of the matrix $\{x_i, A_j\}$ drops. Because the p_i are conjugate to the x_i this matrix of Poisson brackets can also be seen as the Jacobian matrix of the A_j with respect to p_i, $\{x_i, A_j\} = \partial A_j/\partial p_i$. But the rank of this Jacobian drops at the critical points of the A_j over the p_is and hence when the image of the projection onto x_i space has vertical directions above it. These last two conditions are precisely our definitions of the critical set and the caustic set from the last section.

For the 2D oscillator the caustic sets are the two lines on which $p_1 = 0$ ($x_1 = x_{1\min}$ or $x_{1\max}$) and the two lines on which $p_2 = 0$ ($x_2 = x_{2\min}$ or $x_{2\max}$). These lines meet at four corner points where $p_1 = p_2 = 0$. All of this fits together nicely with the previous discussion of caustics; the lines on which $p_1 = 0$ are also the lines at which the two branches $p_1 = \pm\sqrt{2E_1 - \omega_1^2 x_1^2}$ join. Furthermore, the configuration space

[8] A more sophisticated approach is to develop "uniform approximations" that are non-singular across the caustic [41]. Littlejohn and Yu have developed a uniform approximation to the $6j$-symbol, a particularly useful spin network, in [22].

orbit is always tangent to the caustic boundary and so there are only two tangent directions available at these tangencies, this is illustrated in the Lissajous Figure 2.9.

This gives a rough way to picture how the oscillator torus sits above configuration space: It is like a bicycle tire sitting upright over the pavement. The vertical projection collapses the four branches down onto a rectangular patch of pavement and the central hole is not visible in the projected image. This is only a rough picture because the torus is embedded in $\mathbb{R}^4$. In particular, along the caustic lines $p_1 = 0$ and $p_2 = 0$, and solely along these lines, there are just two branches lying above each point, while at the corners where these lines meet there is a single branch. These features cannot be accurately captured with a three dimensional embedding. A great deal can be learned by calculating the amplitude determinant and thinking about the consequences for the caustic and the Lagrangian manifold.

We turn now to consideration of quantization conditions. The classical phase space $\mathbb{R}^4$ of our 2D oscillator is foliated by the toric leaves of constant H and K_3 as discussed in section 2.1. Not all of these tori support wavefunctions. At the level of the semiclassical wavefunction this has a simple interpretation: the Lagrangian manifold must satisfy certain conditions in order for the wavefunction to be single valued. More physically, the selection of the appropriate "quantized" Lagrangian manifolds is how the quantization conditions enter the theory.

At the end of the discussion of Lagrangian manifolds the observation was made that the topology of the Lagrangian manifold could lead to a topological multivaluedness of the action integrals

$$S_k = \int p dq. \tag{2.56}$$

By requiring that this topological proliferation not effect the physics, i.e. the value of the wavefunction, a quantization condition is imposed on the system. Consider the one degree of freedom case. In this case the only topologically non-trivial Lagrangian manifold is a topological circle and for a time independent Hamiltonian this is generally the level set $H = E$. Because the action appears in the phase of the WKB *ansatz*, $\exp\{\frac{i}{\hbar}S\}$, requiring that the wavefunction be single valued amounts to requiring that

$$I(E) \equiv \Delta S \equiv \oint p dq = 2\pi n\hbar = nh. \tag{2.57}$$

This is the famous Bohr-Sommerfeld quantization condition. Geometrically this integral is the phase space area contained within the level set $H = E$. This area only depends on the energy and so we find that $I = I(E)$. Unfortunately, but somewhat reasonably, this function is also called an action.

The period just following the introduction of this quantization condition has a fascinating history that we briefly discuss because of the physical insight that it lends. Soon after Bohr worked out the spectrum of the Hydrogen atom using the sort of result discussed above, Wilson, Sommerfeld and Epstein began to try to understand

how to properly formulate it for systems of more than one degree of freedom. They found that when the Hamilton-Jacobi equation could be treated by separation of variables in a particular coordinate system then one could write down a quantization condition of this sort for each separated coordinate with a periodic evolution. This insight leads into the theory of action-angle variables. A year later, Einstein returned to thinking about these issues after his concentrated work on general relativity during 1911-1915. With coordinate covariance at the front of his mind he was dissatisfied with the reliance of the Sommerfeld result on separation of variables and the choice of a particular coordinate system. Building on Poincaré's discussion at the 1911 Solvay conference that $\sum_i p_i dq^i$ is invariant under canonical (coordinate) transformations Einstein realized that it was the integrable system's tori that were important. A translation of his formulation into modern language including a slight modification yields the quantization conditions,

$$I_j(A) \equiv \oint_{C_j} \sum_{i=1}^{n} p_i dq^i = \left(n_j + \frac{\mu_j}{4}\right) h \qquad j = (1, \ldots, n), \tag{2.58}$$

here the C_j are n independent basis contours (in the sense of homology theory) for the torus and n_j is the quantum number for the jth condition. In less technical terms these are the different circles whose cartesian product makes up the torus, the torus cycles. The μ_j are an additional ingredient as compared to what Einstein worked out and are called Maslov indices. Maslov indices will be discussed at greater length below; for now it is enough to say that for kinetic plus potential problems the Maslov index counts the number of times that the contour C_j crosses a caustic.[9] Once again the actions I_j are functions of the $A = \{A_1, \ldots, A_n\}$ that specify the level set. This is the EBK quantization rule briefly mentioned in the introduction to this section. This rule extends to more complicated topologies, where there are as many quantization conditions as there are independent closed basis contours. Einstein also understood that classically chaotic motions would not give rise to Lagrangian manifolds with simple topologies and that this would raise new difficulties in their quantization. This was more than fifty years before anyone else considered the quantization of classical chaos and its significance.

During this same period Ehrenfest [65] was trying to develop a deeper physical understanding of these conditions. He realized that the loop integral $I(E) = \oint p dq$ is an adiabatic invariant. An adiabatic invariant is a quantity that is invariant as one slowly varies a parameter in the Hamiltonian. Adiabatic invariants allow the quantitative connection of two *different physical systems* when there exists a smooth perturbation connecting their Hamiltonians. This leads to a remarkable conclusion

[9] A subtlety worth noting is that these μ_j are not the same Maslov indices that appear in the wavefunction, although the two types originate from the same phenomenon. For one thing the wavefunction Maslov indices μ_k are labelled by the branch index k while these are labelled by the index j that keeps track of the independent basis contours C_j.

in the case of quantization conditions. As soon as one establishes the quantization of a quantity in one physical system, say the energy of the Hydrogen atom, then that quantity is quantized for all systems that can be adiabatically reached from the first system, e.g. the energies of a Hydrogen atom in an electric field.[10] Quantization is not an isolated result that needs proving for each and every system one at a time.

The 2D oscillator quantization conditions take their simplest form using the cycles specified by $H_1 = E_1$ and $H_2 = E_2$. The first quantization condition arises from integrating $p_1 dq_1$ along the basis contour of constant q_2, call it C_1, and requiring that,

$$I_1 = \oint_{C_1} p_1 dq_1 = \oint_{C_1} \sqrt{2E_1 - \omega_1^2 x_1^2} dx_1 = (n_1 + \frac{1}{2})h. \tag{2.59}$$

To keep the expressions readable it is useful to abbreviate $x_{1\max} = \sqrt{2E_1/\omega_1^2}$ to x_{1m}, then $x_{1\min} = -x_{1m}$ and because both the direction of integration and the sign of the momentum changes at these turning points we have,

$$I_1 = \oint_{C_1} p_1 dq_1 = 2\int_{-x_{1m}}^{x_{1m}} \sqrt{2E_1 - \omega_1^2 x_1^2} dx_1. \tag{2.60}$$

Changing variables from x_1 to θ_1 defined by $x_1 = x_{1m}\sin(\theta_1)$ and with $dx_1 = x_1\cos(\theta_1)d\theta_1$ gives,

$$\begin{aligned} I_1 &= 2\sqrt{2E_1}x_{1m}\int_{\sin^{-1}(-1)}^{\sin^{-1}(1)} \sqrt{1-\sin^2(\theta_1)}\cos(\theta_1)d\theta_1 \\ &= \frac{2E_1}{\omega_1}\left[\theta_1 + \frac{1}{2}\sin^2\theta_1\right|_{\sin^{-1}(-1)}^{\sin^{-1}(1)} = \frac{2\pi E_1}{\omega_1}. \end{aligned} \tag{2.61}$$

This yields the exact quantization condition for the harmonic oscillator, $E_1 = (n_1 + 1/2)\hbar\omega_1$. The same calculation leads to $E_2 = (n_2 + \frac{1}{2})\hbar\omega_2$. Combining these two results yields the quantization of the total energy and the energy difference,

$$E_{n_1,n_2} = (n_1 + \frac{1}{2})\hbar\omega_1 + (n_2 + \frac{1}{2})\hbar\omega_2 \tag{2.62}$$

$$K_{3n_1,n_2} = (n_1 + \frac{1}{2})\hbar\omega_1 - (n_2 + \frac{1}{2})\hbar\omega_2. \tag{2.63}$$

The calculation of the phases for the 2D oscillator is a mild extension of the calculation here for the quantization condition. Recall that the four branches above (x_{10}, x_{20}) can be distinguished by the sign of the momenta on that branch, $\{(p_{10}, p_{20})$, $(-p_{10}, -p_{20})$, $(-p_{10}, p_{20})$, $(p_{10}, -p_{20})\}$, we label these branches with the abbreviated label set $k \in \{a, b, c, d\}$, so that S_a refers to the action on the branch containing (p_{10}, p_{20}), and so on. In order to evaluate the action integrals S_k it is necessary to

[10] Two interesting modern treatments of the Stark effect using semiclassical means are in [66, 67].

choose a conventional starting point on the Lagrangian manifold. This amounts to an overall phase convention for the wavefunction. All four branches of the 2D oscillator coalesce at the 4 extreme boundary points of the caustic rectangle (see Fig. 2.9). We arbitrarily choose $(x_{1\min}, x_{2\min})$ to be the conventional starting point. For this problem there is a natural choice of contour γ along which to evaluate $S_k = \int \sum pdq$: starting at $(x_{1\min}, x_{2\min})$ integrate along the "straight line" contour γ_1 with $x_2 = x_{2\min}$ until x_{10} is reached, then integrate along the "straight line" contour γ_2 with $x_1 = x_{10}$ until x_{20} is reached, so $\gamma = \gamma_1 + \gamma_2$. These contours are straight lines in the configuration space and, of course, the integrals are performed along the corresponding lifted curve in the branch under current consideration. Each individual branch is topologically trivial, hence S_k is independent of path and this choice is only natural in the sense that it simplifies calculations. Along γ_1, $p_2 = 0$ and so $\int p_2 dx_2$ doesn't contribute, while along γ_2, $x_1 = x_{10}$ and so $\int p_1 dx_1$ doesn't contribute. Focusing on the branch $k = a$ containing (p_{10}, p_{20}) we find,

$$S_a = \int_{\gamma_1} p_1 dx_1 + \int_{\gamma_2} p_2 dx_2 = \int_{x_{1\min}}^{x_{10}} p_1(x_1) dx_1 + \int_{x_{2\min}}^{x_{20}} p_2(x_2) dx_2. \tag{2.64}$$

Except for the limits of integration these are the same integrals as those performed in the quantization conditions and yield,

$$\begin{aligned} S_a = \frac{E_1}{\omega_1} &\left[\sin^{-1}\left(\frac{x_1}{x_{1m}}\right) + \frac{x_1}{x_{1m}}\sqrt{1-\left(\frac{x_1}{x_{1m}}\right)^2} - \frac{\pi}{2} \right] \\ &+ \frac{E_2}{\omega_2}\left[\sin^{-1}\left(\frac{x_2}{x_{2m}}\right) + \frac{x_2}{x_{2m}}\sqrt{1-\left(\frac{x_2}{x_{2m}}\right)^2} - \frac{\pi}{2} \right] \\ &\equiv S_1 + S_2. \end{aligned} \tag{2.65}$$

Again we adopt the abbreviations $x_{1m} = x_{1\max} = -x_{1\min}$ and $x_{2m} = x_{2\max}$ and we drop the overly explicit (x_{10}, x_{20}) label of a point below the branch in favor of (x_1, x_2). On the other branches all that changes is the sign of the momentum in the action integral and so combining these with the correct Maslov indices (see the next section for the calculation of the Maslov indices) yields,

$$\begin{aligned} S_b - \pi &= -S_1 - S_2 - \pi = -S_a - \pi \\ S_c - \frac{\pi}{2} &= S_1 - S_2 - \frac{\pi}{2} \\ S_d - \frac{\pi}{2} &= -S_1 + S_2 - \frac{\pi}{2} = -S_c - \frac{\pi}{2}. \end{aligned} \tag{2.66}$$

These calculations are straightforward but now we can use them to illustrate an important technical aspect of the general semiclassical result for the wavefunction (2.48). When the physical system under consideration has a discrete symmetry this

can be leveraged to collapse the sum on branches. In the present example the symmetry is time reversal and it leaves the spatial coordinates invariant, $(x_1, x_2) \mapsto (x_1, x_2)$, while reversing the signs of the momenta $(p_1, p_2) \mapsto (-p_1, -p_2)$. The importance of time reversal is that it ties the different branches of the Lagrangian manifold together. It is because of time reversal that the intersection points $(-p_{10}, -p_{20})$ and (p_{10}, p_{20}) only differ by signs (here the shared x dependence has been repressed). From this and the absolute values in their definition it follows that the amplitude factors R_k for all four branches are the same,

$$R_k = \frac{1}{\sqrt{2|p_{10}p_{20}|}}, \tag{2.67}$$

and can be factored out of the sum (2.48). Time reversal also has a dramatic simplifying effect on the phases. This is best illustrated by direct computation, which we turn to now.

Using equations (2.65) and (2.66) the sum of the phases, actions plus Maslov corrections, can be written as,

$$\begin{aligned}
\sum_{k=a}^{d} e^{i(S_k - \mu_k \pi/2)} &= e^{iS_a} + e^{i(S_b - \pi)} + e^{i(S_c - \pi/2)} + e^{i(S_d - \pi/2)} \\
&= e^{-i\pi/2}\left(e^{i(S_a + \pi/2)} + e^{i(-S_a - \pi/2)} + e^{iS_c} + e^{-iS_c}\right) \\
&= -i\left(2\cos\left(S_a + \frac{\pi}{2}\right) + 2\cos\left(S_c\right)\right) \\
&= -i4\left(\cos\left(S_1 + \frac{\pi}{4}\right)\cos\left(S_2 + \frac{\pi}{4}\right)\right).
\end{aligned} \tag{2.68}$$

For notational clarity we have set $\hbar = 1$ in these expressions. The collapse of the sum of exponentials is characteristic of problems with a discrete symmetry. The factor $-i$ in front of (2.68) cancels with the i outside the sum in (2.48). This cancellation was engineered. The overall phase depends on the initial point used to calculate the action integrals (2.65) and (2.66) and is completely conventional. This corresponds to the usual arbitrariness of the overall phase in quantum mechanics.

The details recorded, assembling the semiclassical wavefunction is now simple: $V_A = 2\pi^2/(\omega_1\omega_2)$, $R_k = 1/\sqrt{2|p_1p_2|}$ and

$$\psi(x_1, x_2) = \frac{2}{\pi}\sqrt{\frac{\omega_1\omega_2}{|p_1p_2|}}\cos\left(S_1 + \frac{\pi}{4}\right)\cos\left(S_2 + \frac{\pi}{4}\right), \tag{2.69}$$

with

$$S_1 = \frac{E_1}{\omega_1}\left[\sin^{-1}\left(\frac{x_1}{x_{1m}}\right) + \frac{x_1}{x_{1m}}\sqrt{1 - \left(\frac{x_1}{x_{1m}}\right)^2} - \frac{\pi}{2}\right], \tag{2.70}$$

$E_1 = (n_1 + 1/2)\hbar\omega_1$ and $x_{1m} = \sqrt{2E_1/\omega_1^2}$; the action S_2, energy E_2 and constant x_{2m} are given by the same formulae with the subscript 1 replaced by the subscript 2. The exact harmonic oscillator wavefunction is

$$\psi(x_1, x_2) = \left(\frac{\omega_1\omega_2}{\pi^2\hbar^2}\right)^{1/4} \frac{1}{\sqrt{2^{n_1 n_2} n_1! n_2!}} H_{n_1}(\chi_1)\, H_{n_2}(\chi_2)\, e^{-\chi_1^2/2 - \chi_2^2/2}, \tag{2.71}$$

where $\chi_1 = \sqrt{\omega_1/\hbar} \cdot x_1$ and similarly for χ_2. Here $H_n(\chi)$ is the Hermite polynomial. Casting the Hermite polynomials as integrals and performing a stationary phase approximation shows that these two solutions agree in the asymptotic limit $n_1 \to \infty, n_2 \to \infty$. We forgo this analytic comparison in favor of a direct numerical one that shows that the approximation is excellent in the classically allowed region even for small n_1, n_2, see Figure 2.10. Of course, classically forbidden regions can also

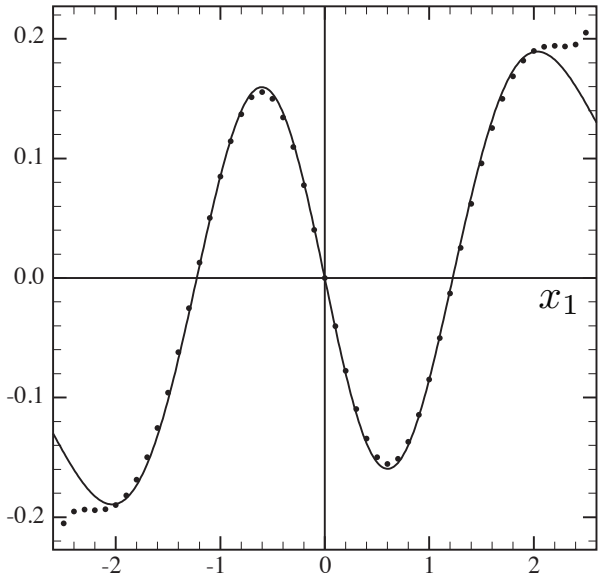


Figure 2.10: A comparison of the exact (solid line) and WKB (dots) wavefunctions for the 2D oscillator. The plot is along a constant $x_2 = -1$ cross-section with $n_1 = 3$ and $n_2 = 2$.

be treated semiclassically, see [41]. This illustrates the accuracy of these asymptotic approximations discussed in the introduction to this section.

Before proceeding to a discussion of the calculation of Maslov indices we summarize the general formulation of the semiclassical wavefunction. The semiclassical wavefunction is given by the formula,

$$\langle b|a\rangle = \frac{(2\pi i)^n}{V_A V_B} \sum_k R_k \exp\left\{i\left[S_A(\alpha_k) - S_B(\beta_k) - \mu_k\pi/2\right]\right\}. \tag{2.72}$$

This is an approximation to the wave function determined by the states of the complete sets of commuting operators $\{\hat{A}_1, \ldots, \hat{A}_n\}$ and $\{\hat{B}_1, \ldots, \hat{B}_n\}$. Corresponding to these operators are the principle symbols $A = \{A_i\}_{i=1,\cdots n}$ and $B = \{B_i\}_{i=1,\cdots n}$, whose level set values a and b determine two quantized manifolds. The overlap of these states is calculated by stationary phase integration in the x-representation and the stationary phase points are given by the intersections of the Lagrangian manifolds determined by a and b. The sets of angles conjugate to A and B are denoted by α and β with their values at the stationary phase point on branch k denoted α_k and β_k, (each of these values determines the same point in Φ). The action S_A is determined by integrating from a conventional starting point along the A-manifold and similarly for S_B and the volumes V_A and V_B are the volumes of each of these manifolds measured using the volume forms determined by α and β respectively. Finally, μ_k is the Maslov index for the kth branch and the amplitude can be expressed in terms of the amplitude determinant,

$$R_k = |D_k|^{-1/2} = |\det\{A_i, B_j\}|^{-1/2}. \tag{2.73}$$

The formula (2.72) gives the semiclassical wavefunction in the generic case where the Lagrangian manifolds determined by a and b intersect in discrete points. In Chapter 5 the semiclassical treatment of the $9j$-symbol will lead to Lagrangian manifolds with non-generic, higher dimensional intersections. In [63] Littlejohn also treats this case. Let us assume that the intersection is $(n-r)$-dimensional and that the functional dependence of the B observables on the A observables can be made explicit so that $B = \{B_1, \ldots, B_r, A_{r+1}, \ldots, A_n\}$. The stationary phase points are again given by the ($(n-r)$-dimensional) intersections of the two Lagrangian manifolds and the wavefunction is,

$$\langle b|a\rangle = \frac{(2\pi i)^r}{\sqrt{V_A V_B}} \sum_k V_k R_k \exp\{i\,[S_A(\alpha_k) - S_B(\beta_k) - \mu_k \pi/2]\}, \tag{2.74}$$

where V_k is the volume of the kth intersection (with respect to the form $\alpha_{r+1} \wedge \cdots \wedge \alpha_n$). The amplitude determinant is given by equation (2.73) but the i and j indices are understood to only run over the functionally independent observables, $i, j = 1, \ldots, r$. The action difference $S_A(\alpha_k) - S_B(\beta_k)$ does not depend on what point of the intersection it is evaluated at because the integral of pdx back and forth along a path lying in the intersection is zero. This is the most general formulation of the semiclassical wavefunction that will be needed in this dissertation. The two formulae (2.72) and (2.74) encapsulate the geometrical approach to quantum wavefunctions furnished by semiclassical methods. We return now to a more complete treatment of Maslov indices.

In the above calculation of the 2D oscillator wave function the Maslov indices were inserted by hand. This is not uncommon in practice. The calculation of Maslov indices can be cumbersome and in simpler examples experience can guide a good guess

that can be checked numerically. The appropriate Maslov indices for the asymptotics of the $9j$-symbol, presented in Chapter 5, were difficult to find. This suggests that a better tool will need to be developed for practical purposes in finding the asymptotics of higher $3nj$-symbols. In the next section we present an approach developed by Robert Littlejohn in unpublished work; this seems a promising approach to us.

The reader new to semiclassics is warned that Maslov indices are a more advanced topic, the next section can be safely skipped without significant cost to the big picture. The qualitative picture for Maslov indices is simple to state: Just as the Bohr-Sommerfeld condition ensures that the action integral around a closed curve gives a consistent value, $S_0 \to S_0 + 2\pi n\hbar$, the Maslov index ensures that the action integrals on different branches are consistent with one another when one extends the integration through a caustic. The next section develops this qualitative picture in more detail.

2.2.4 Maslov indices

The Maslov index allows the consistent comparison of the actions developed on two different branches separated by a caustic. In general the action $S = \int p(q)dq$ is not defined at a caustic, for example p may not be a function of q. The key idea to overcoming this difficulty is to switch representation spaces, for example going to the momentum representation in the neighborhood of the caustic. This idea was introduced by Maslov [68, 69]. Once in the momentum representation the integral can be carried through the q-space caustic because the singular set can have an ordinary projection onto the momentum space. On the other side of the caustic returning to the q-representation yields an expression for the action on the new branch.

Consider the one degree of freedom example depicted in Figure 2.11. The labels 1 and 2 will be used both for the upper and lower branches and also for the points 1 and 2 on the upper and lower branches. Recall the form of the wavefunction,

$$\psi(x) = \langle x|a\rangle = \sum_k R_k e^{i(S_k - \mu_k \pi/2)}, \tag{2.75}$$

where the action is $S = \int_L pdx$, with L the Lagrangian manifold $A = a$. We set $\hbar = 1$ for the remainder. The first order of business here is to briefly sketch the origin of the factors μ_k.

Some notation will simplify what follows. The difference in phase between branch 1 and branch 2 is,

$$\begin{aligned} \Delta\text{phase} &= \text{phase}_2 - \text{phase}_1 \\ &= S_2 - S_1 - (\mu_2 - \mu_1)\frac{\pi}{2} \\ &= \Delta S - (\Delta\mu)\frac{\pi}{2}, \end{aligned} \tag{2.76}$$

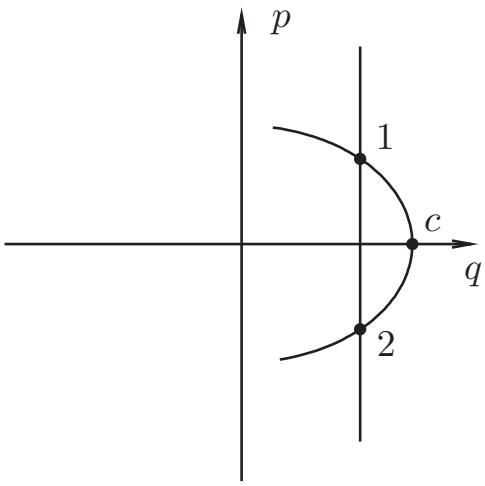


Figure 2.11: A Lagrangian manifold defined by $A = a$, the caustic point c is marked with a dot. The numbers 1 and 2 label both the upper and lower branches and the intersection points 1 and 2 respectively.

where the last line defines $\Delta S = \int_{1L}^{2} p dx$ and $\Delta\mu = \mu_2 - \mu_1$ the change in Maslov index. It is not hard to show that the momentum space wavefunction can be consistently approximated with a similar WKB *ansatz* $\phi(p) \sim \exp\{iT(p)\}$. Here the momentum space action $T(p)$, is given by

$$T(p) = -\int_L x dp, \tag{2.77}$$

and $x = x_L(p)$, is the position field determined by the momentum action $T(p)$ corresponding to the Lagrangian manifold L. The derivatives of the momentum action are,

$$T'(p) = -x = -x_L(p), \tag{2.78}$$

$$T''(p) = -x'_L(p) = -\left(\frac{\partial x}{\partial p}\right)_A. \tag{2.79}$$

In the second equation the A subscript indicates that the partial derivative is taken at constant A. Note that the sign of T'' is minus the sign of the slope of the tangent to L, that is, $\mathrm{sgn}(T'') = -\mathrm{sgn}(dp/dx)$. Suppose that the momentum space wave function $\phi(p)$ corresponding to the L of Figure 2.11 is given. We are interested in the inverse Fourier transform that specifies this wavefunction in the q-representation. This Fourier transform can be evaluated by stationary phase, an approximation at the same order as the WKB treatment being studied here. It turns out that the amplitude

is continuous across the caustic and so we focus only on the phase contribution,

$$\int dp e^{i/2\{T''(p_c)(p-p_c)\}} = \sqrt{\frac{2\pi}{|T''|}} \begin{cases} e^{i\pi/4} & \text{if} \quad T'' > 0 \\ e^{-i\pi/4} & \text{if} \quad T'' < 0. \end{cases} \tag{2.80}$$

The two cases on the right side of the equation are the source of the Maslov index. At the caustic $T''(p_c) = 0$, and the sign of T'' on either side of this zero crossing determines $\Delta\mu$ by,

$$T'' \quad \text{goes} \quad + \to - \implies \Delta\mu = +1, \tag{2.81}$$

$$T'' \quad \text{goes} \quad - \to + \implies \Delta\mu = -1. \tag{2.82}$$

Geometrically if the tangent line with slope dp/dx goes from $\text{sgn}(dp/dx) = -1$ to $\text{sgn}(dp/dx) = +1$ then $\Delta\mu = +1$ and conversely. In the example above, going from branch 1 to branch 2, this amounts to $\Delta\mu = +1$ and so if we conventionally set $\mu_1 = 0$ then $\mu_2 = +1$.

Now that we've sketched where the Maslov indices originate we move to the true focus of this section, their calculation. First we will generalize to arbitrary one degree of freedom wavefunctions expressed in general bases. As before we consider the A Lagrangian manifold specified by $A = a$. Take the basis to be specified by a function B, whose level sets determine a foliation of phase space into curves $B = b$. First a technical point: the B is like an action function, in the sense of action-angle variables, or generalized momentum, so specifying our basis by $B = b$ involves a change of sign. This is because the canonical transformation that switches q's and p's involves a sign, $\tilde{q} = p$ and $\tilde{p} = -q$, with $(\tilde{q}, \tilde{p})$ the new coordinates. Thus, the correspondence is $x \to b$ (or B) and $p \to -\beta$, where β is the conjugate angle to B. This means that, introducing the shorthand $\lambda = T''$,

$$\lambda = T'' = + \left(\frac{\partial b}{\partial \beta}\right)_A. \tag{2.83}$$

Let the Hamiltonian flow of A, X_A be denoted by an overdot $X_A f = \partial f/\partial\alpha = \dot{f}$. Then,

$$\lambda = \left(\frac{\partial b}{\partial \beta}\right)_A = \left(\frac{\partial b}{\partial \alpha}\right)_A \left(\frac{\partial \alpha}{\partial \beta}\right)_A = \frac{\dot{b}}{\dot{\beta}} = \frac{\{B, A\}}{\{\beta, A\}}. \tag{2.84}$$

The change in sign of T'' at the caustic is characterized by $\dot{\lambda}|_c = X_A T''|_c$. So we calculate,

$$\dot{\lambda} = \frac{\ddot{b}}{\dot{\beta}} - \frac{\dot{b}\ddot{\beta}}{\dot{\beta}^2}, \tag{2.85}$$

and at the caustic, $\dot{b} = \{B, A\} = 0$, this simplifies to,

$$\dot{\lambda}|_c = \frac{\ddot{b}}{\dot{\beta}} = \frac{\{\{B, A\}, A\}}{\{\beta, A\}}. \tag{2.86}$$

The denominator $\dot{\beta} = \{\beta, A\}$ can be difficult to calculate directly, interestingly this can be avoided at the caustic. The caustic condition can be expressed in terms of the Poisson bivector $\Pi \in \Lambda^2\Phi$ or in terms of the symplectic form $\omega \in \Omega^2\Phi$,

$$\{B, A\} = 0 = \Pi(dB, dA) = -\omega(X_B, X_A). \tag{2.87}$$

The vanishing of the symplectic form implies that the vectors X_A and X_B are proportional here, call the proportionality factor e so that $X_A = eX_B$ or because $\Pi(dB, dA) = 0$,

$$dA = edB. \tag{2.88}$$

This means that at the caustic,

$$\dot{\beta} = \{\beta, A\} = \left(\frac{\partial A}{\partial B}\right)_\beta = e. \tag{2.89}$$

Then, in one degree of freedom the algorithm is (1) find the caustic set, $\{A, B\} = 0$ (2) find e from $dA = edB$ at the caustic and (3) calculate,

$$\dot{\lambda}|_c = \frac{1}{e}\{\{B, A\}, A\}. \tag{2.90}$$

We illustrate this algorithm with a kinetic plus potential Hamiltonian $A = H = p^2/2 + V(x)$. In the x-representation the caustic set is determined by $\{x, H\} = p = 0$ and we traverse it along the dynamics so that,

$$\{\{B, A\}, A\} = \{\{x, H\}, H\} = \{p, H\} = -V'(x). \tag{2.91}$$

Now, $dB = dx$ and $dA =$dH$= pdp + V'(x)dx$, the latter becoming $dH = V'(x)dx = edx$ at the caustic. So,

$$\dot{\lambda}|_c = \frac{-V'(x)}{V'(x)} = -1, \tag{2.92}$$

and this implies, returning to (2.81), that $\Delta\mu = +1$.

This analysis can be extended to more degrees of freedom. In the one degree of freedom case the Maslov index was determined by the change in sign of T'' as the caustic was traversed. In more degrees of freedom the Maslov index is given by the number of eigenvalues of $\partial^2 T/\partial p_i \partial p_j$ that change sign as one traverses the caustic. Once again we will calculate this by finding the derivatives of these eigenvalues at the caustic.

The momentum space action now depends on all of the momenta p_i $(i = 1, \ldots, n)$ and its derivatives are,

$$\frac{\partial T}{\partial p_i} = -x_i = -x_{Li}(p), \tag{2.93}$$

$$\frac{\partial^2 T}{\partial p_i \partial p_j} = -\left(\frac{\partial x_i}{\partial p_j}\right). \tag{2.94}$$

The extension to $A = \{A_1, \ldots, A_n\}$ and $B = \{B_1, \ldots, B_n\}$ observables gives,

$$T_{ij} \equiv + \left(\frac{\partial b_i}{\partial \beta_j} \right)_A = \left(\frac{\partial b_i}{\partial \alpha_k} \right)_A \left(\frac{\partial \alpha_k}{\partial \beta_j} \right)_A . \tag{2.95}$$

This factorization leads once again to the introduction of

$$D_{ik} \equiv \left(\frac{\partial b_i}{\partial \alpha_k} \right)_A = \{B_i, A_k\}, \tag{2.96}$$

the matrix of Poisson brackets that gives rise to the amplitude determinant $D = \det D_{ik}$. The second factor in (2.95), or rather its inverse will be denoted by

$$E_{jk} \equiv \left(\frac{\partial \beta_j}{\partial \alpha_k} \right)_A \tag{2.97}$$

and satisfies

$$E_{jk} E^{-1}_{kl} = \left(\frac{\partial \beta_j}{\partial \alpha_k} \right)_A \left(\frac{\partial \alpha_k}{\partial \beta_j} \right)_A = \delta_{jl}. \tag{2.98}$$

With these definitions,

$$T_{ij} = D_{ik} E^{-1}_{kj} = T_{ji}. \tag{2.99}$$

As in the one degree of freedom case the caustic set separates the Lagrangian manifold's branches. In order to specify the transition from one branch to the other we choose a curve γ that is transversal to the caustic. For the remainder of this section an overdot indicates the derivative with respect to the parameter τ along the curve γ. We will be developing an algorithm to evaluate the change in Maslov index as one traverses the caustic along this curve.

Now, consider a normalized eigenvector e_i of T_{ij}, with eigenvalue λ. Because T_{ij} is symmetric there is no need to distinguish left and right eigenvectors, however it will be convenient to think of this as a left eigenvector,

$$e_i T_{ij} = \lambda e_j. \tag{2.100}$$

Suppose that $\lambda = 0$ at the caustic, we are interested in $\dot{\lambda} = d\lambda/d\tau$. The Feynman-Hellman theorem guarantees that

$$\dot{\lambda} = e_i \dot{T}_{ij} e_j. \tag{2.101}$$

Using the definitions of D_{ij} and E_{ij} introduced above, $T_{ij} = D_{ik} E^{-1}_{kj}$ and we have,

$$\dot{T}_{ij} = \dot{D}_{ik} E^{-1}_{kj} - D_{il} E^{-1}_{lm} \dot{E}_{mn} E^{-1}_{nj}. \tag{2.102}$$

Putting this into (2.101) yields,

$$\dot{\lambda} = e_i \dot{T}_{ij} e_j = e_i \dot{D}_{ij} E^{-1}_{jk} e_k - e_i T_{ij} \dot{E}_{jk} E^{-1}_{kl} e_l. \tag{2.103}$$

This holds in general while at the caustic $\lambda = 0$ so that $e_i T_{ij} = 0$ and

$$\dot{\lambda}|_c = e_i \dot{D}_{ij} E^{-1}_{jk} e_k. \tag{2.104}$$

As above we would like to avoid computing E^{-1}_{ij} directly. Instead note that the rank of $D_{ij} = \{B_i, A_j\}$ drops at the caustic (in general that of E^{-1}_{ij} does not) and so,

$$e_i T_{ij} = 0 = e_i D_{ik} E^{-1}_{kj}. \tag{2.105}$$

This implies that $e_i D_{ik} = 0$ at the caustic. Writing this condition out,

$$e_i \{B_i, A_j\} = 0 \tag{2.106}$$

or once again,

$$e_i \Pi(dB_i, dA_j) = -e_i \omega(X_{B_i}, X_{A_j}) = -\omega(e_i X_{B_i}, X_{A_j}) = 0. \tag{2.107}$$

This hold for all j and so the non-degeneracy of ω implies that $\text{span}\{X_{A_j}\} = \text{span}\{X_{B_j}\}$. But then the vector $Y = e_i X_{B_i}$ can be expanded, with coefficients f_j say, in terms of the X_{A_j},

$$Y = e_i X_{B_i} = f_j X_{A_j} \tag{2.108}$$

or what is the same,

$$e_i dB_i = f_j dA_j. \tag{2.109}$$

This important identity (the analog of (2.88)) allows the elimination of E^{-1}_{ij} in favor of f_j:

$$\begin{aligned} E_{ij} f_j &= \{\beta_i, A_j\} f_j \\ &= -\omega(X_{\beta_i}, X_{A_j} f_j) \\ &= -\omega(X_{\beta_i}, e_j X_{B_j}) \\ &= \{\beta_i, B_j\} e_j = \delta_{ij} e_j = e_i. \end{aligned} \tag{2.110}$$

This holds at the caustic and there it implies $f_i = E^{-1}_{ij} e_j$. Putting this into (2.104) yields,

$$\dot{\lambda}|_c = e_i \dot{D}_{ij} E^{-1}_{jk} e_k = e_i \dot{D}_{ij} f_j. \tag{2.111}$$

Then the algorithm for multiple degrees of freedom is: (1) Find the caustic set $D = \det\{B_i, A_j\} = 0$. (2) Find a left null eigenvector e_i of D_{ij} evaluated on the caustic. (3) Find the corresponding f_j such that $e_i dB_i = f_j dA_j$ and (4) evaluate $\dot{\lambda}|_c$ via

$$\dot{\lambda}|_c = e_i \dot{D}_{ij} f_j, \tag{2.112}$$

where $\dot{D}_{ij} = d/d\tau \{B_i, A_j\} = \{\{B_i, A_j\}, C\}$ and here C generates the flow along γ. In practice it can be convenient to take $C = A_k$ (for some fixed k).

We conclude this section and illustrate the algorithm above by performing the promised calculation of the Maslov indices for the 2D oscillator. The matrix of Poisson brackets was calculated above (2.51),

$$D_{ij} = \begin{pmatrix} p_1 & p_1 \\ p_2 & -p_2 \end{pmatrix} \tag{2.113}$$

and has determinant $D = -2p_1p_2$. The critical sets are given by $p_1 = 0$ or $p_2 = 0$ and project onto the caustic sets illustrated by the dashed line in Figure (2.9). The formalism we have developed here can be applied to find the Maslov index as one crosses any of the configuration space lines on which $p_1 = 0$ or $p_2 = 0$.[11] Focusing on the case with singular set $p_1 = 0$ and caustic projection along the line $x_1 = x_{1\min}$ consider traversing from the branch containing $(+p_{10}, +p_{20})$ to that containing $(-p_{10}, p_{20})$. To implement this we choose a dynamical orbit γ and hence the overdot indicates a time derivative and the flow is generated by H. At the caustic the matrix of Poisson brackets becomes

$$D_{ij} = \begin{pmatrix} 0 & 0 \\ p_2 & -p_2 \end{pmatrix}. \tag{2.114}$$

A left null eigenvector of this matrix is $(e_1, e_2) = (1, 0)$. The definition of the f_j, $e_i dB_i = f_j dA_j$ leads to

$$\begin{aligned} dx_1 &= f_1 dH + f_2 dK_3 \\ &= (f_1 + f_2)\omega_1^2 x_1 dx_1 + (f_1 - f_2)p_2 dp_2 + (f_1 - f_2)\omega_2^2 x_2 dx_2, \end{aligned} \tag{2.115}$$

where in the last equality the caustic condition $p_1 = 0$ has been used. Requiring the vanishing of the coefficients of dx_1 and dp_2 gives the system,

$$\begin{aligned} (f_1 + f_2) &= \frac{1}{\omega_1^2 x_1}, \\ f_1 - f_2 &= 0, \end{aligned} \tag{2.116}$$

with solution $f_1 = f_2 = 1/(2\omega_1^2 x_1)$. The time derivative of D_{ij} is

$$\dot{D}_{ij} = \begin{pmatrix} \{p_1, H\} & \{p_1, H\} \\ \{p_2, H\} & \{-p_2, H\} \end{pmatrix} = \begin{pmatrix} -x_1 & -x_1 \\ -x_2 & x_2 \end{pmatrix} \tag{2.117}$$

and so at the caustic,

$$\dot{\lambda}_1|_c = e_i \dot{D}_{ij} f_j = (1\ 0) \begin{pmatrix} -x_1 & -x_1 \\ -x_2 & x_2 \end{pmatrix} \begin{pmatrix} \frac{1}{2\omega_1^2 x_1} \\ \frac{1}{2\omega_1^2 x_1} \end{pmatrix} = -\frac{1}{\omega_1^2}. \tag{2.118}$$

[11] Note, however, that it cannot be applied at the quadratically singular points $p_1 = p_2 = 0$ because these points are not well described locally by the tangent space. Presumably a higher jet formulation is necessary at these points. This limitation is not serious, passing through the quadratic singular point can be replaced by sequential traversals of the usual singular lines.

This leads to a change in Maslov index between the a- and c-branches, following the dynamics, of $\Delta\mu_{a\to c} = +1$.[12] Taking the convention that the branch containing $(+p_{10}, +p_{20})$ has Maslov index 0 then $\mu_a = 0$ and $\mu_c = 1$. Recall the index convention of the last section $\{a, b, c, d\} = \{(p_{10}, p_{20}), (-p_{10}, -p_{20}), (-p_{10}, p_{20}), (p_{10}, -p_{20})\}$. A very similar calculation leads to $\dot{\lambda}_2|_c = -1/\omega_2^2$, where λ_2 is the eigenvalue that goes through zero across a caustic determined by $p_2 = 0$. Again the caustic has been traversed along the flow determined by the dynamics of the oscillator. Then, once again with the convention that $\mu_a = 0$, we have $\mu_b = +2$ and $\mu_d = 1$ (the former arises from crossing each caustic once). These are the Maslov indices used in (2.66). A schematic summary of the Maslov indices is provided in Figure 2.12.

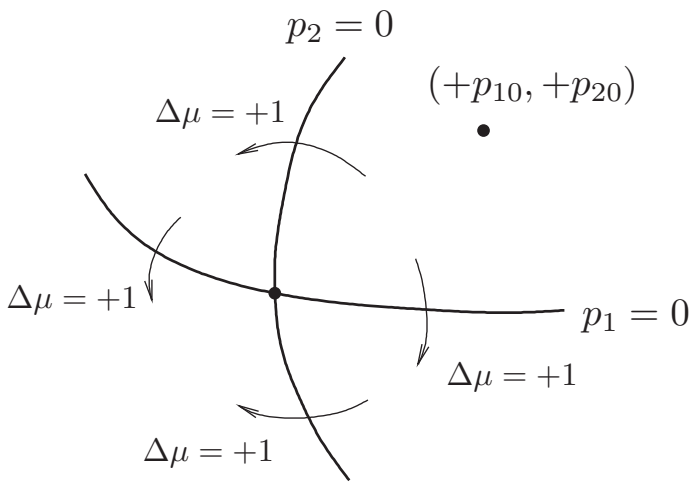


Figure 2.12: Maslov indices for the 2D oscillator.

2.3 Schwinger's model for angular momentum and the Hopf map

In our work on spin networks [37, 24, 70, 71], we have found it invaluable to situate angular momentum theory in different spaces. These spaces have different dimensions and kinematics (that is symplectic and poisson forms) and the calculation

[12] The c subscript is overloaded here, $\lambda|_c$ means the eigenvalue at the caustic, while $\Delta\mu_{a\to c}$ means the change in Maslov index going from branch a to branch c.

of each aspect of the semiclassical wavefunctions (amplitude, phase, etc.) is easiest in a different space. Despite their differences these spaces are connected one to another. These connections are variations on the themes of symplectic and Poisson reduction [57, 58, 72]. In this section we once more take an example-centered approach and illustrate all of this in the simplest possible context of a single angular momentum. Similar procedures apply to collections of angular momenta, [37, 24, 70]. All of the geometry in this thesis is based on the dual Lie algebra $\mathfrak{su}(2)^*$ of $SU(2)$. This Lie algebra is isomorphic to its dual $\mathfrak{su}(2)$ and both are isomorphic to $\mathbb{R}^3$. For technical reasons having to do with momentum maps we focus on the dual algebra $\mathfrak{su}(2)^*$. In the context of classical physics this is naturally regarded as the vector space $\mathbb{R}^3$ of angular momenta $\boldsymbol{J}$. Although our applications will overlay new physical interpretations on these vectors, all of the language and intuition of angular momenta applies.

The first space we will consider was introduced by Schwinger [73] and has been used in numerous studies, e.g. [74, 75, 76]. This is the phase spae $\mathcal{S} \equiv \mathbb{C}^2 \cong \mathbb{R}^4$ of a two dimensional harmonic oscillator with $\omega_1 = \omega_2 = 1$, exactly the space that we have considered extensively above.[13] We briefly recall several definitions introduced in section 2.1: the coordinates (x_1, p_1, x_2, p_2) of $\mathbb{R}^4$ correspond to a pair of complex coordinates (z_1, z_2) of $\mathbb{C}^2$ defined by $z_A = (x_A + ip_A)/\sqrt{2},\ A = 1, 2$. Here and in what follows we will use uppercase Latin indices for two component spinors and occasionally we will also suppress spinor indices altogether so that $z = (z_1, z_2)$.[14] The 2D oscillator has four constants of the motion

$$\begin{aligned} H &= \frac{1}{2}(p_1^2 + p_2^2 + x_1^2 + x_2^2), & K_1 &= \frac{1}{2}(x_1 x_2 + p_1 p_2), \\ K_2 &= (x_1 p_2 - x_2 p_1), & K_3 &= \frac{1}{2}(p_1^2 + x_1^2) - \frac{1}{2}(p_2^2 + x_2^2), \end{aligned} \tag{2.119}$$

that necessarily satisfy one functional dependency,

$$H^2 = K_1^2 + K_2^2 + K_3^2. \tag{2.120}$$

Schwinger exploited the fact that the three Ks satisfy an $\mathfrak{su}(2)$ algebra with respect to the Poisson bracket. It is convenient to renormalize these generators and we will rename them at the same time to emphasize the relationship with $\mathfrak{su}(2)$. Let

$$I = \frac{1}{2}\bar{z}z = \frac{1}{2}\bar{z}_A z_A = \frac{1}{2}H, \tag{2.121}$$

$$J_i = \frac{1}{2}\bar{z}\sigma^i z = \frac{1}{2}\bar{z}_A \sigma^i_{AB} z_B = \frac{1}{2}K_i. \tag{2.122}$$

[13] The $\mathcal{S}$ here is a mnemonic both for Schwinger and for spinor.

[14] Unfortunately, the convention on indices differs from that in our papers [37, 24], however greek indices will be reserved for tensors referred to a coordinate basis and the convention adopted here is common in the spinor and twistor literature [77, 78].

Then using the Poisson brackets of $\mathcal{S}$ it is easy to check that

$$\{I, J_i\} = 0 \qquad \text{and} \qquad \{J_i, J_k\} = \epsilon_{ijk} J_k, \tag{2.123}$$

the latter defining the $\mathfrak{su}(2)^*$ algebra. Once again the relation $I^2 = \boldsymbol{J}^2 = J_1^2 + J_2^2 + J_3^2$ holds.

There is a projection $\pi : \mathcal{S} = \mathbb{C}^2 \to \mathbb{R}^3$ that produces a new reduced Poisson manifold from $\mathcal{S}$. A Poisson manifold is a manifold with a Poisson bracket defined on it. This projection, our first example of reduction, proceeds in the simplest possible manner. The J_i are constants of the motion, $\{J_i, I\} = 0$, under the flow generated by I and so they do not change along the oscillator orbits. Declare the reduced space to be the quotient of the original space by these orbits, that is to say, identify all the points along the orbit. Then the Poisson bracket on the reduced space is defined as the value of the original Poisson bracket anywhere along the orbit. In coordinates I generates a $U(1)$ action on $(z_1, z_2) = (r_1 e^{i\xi_1}, r_2 e^{i\xi_2})$ that is just multiplication by the phase $e^{i\psi/2}$, where ψ is the conjugate angle to I (this can be seen by solving Hamilton's equations). As is clear from their definition, (2.122), the J_i are invariant under this action and hence can be taken as coordinates of the reduced space $\mathbb{R}^3 \equiv \Lambda$. The reduced space Λ will be called angular momentum space. The Poisson brackets of the coordinate functions (2.123) extend to any functions invariant under the $U(1)$ action generated by I, i.e. to any functions $F(\boldsymbol{J})$ and $G(\boldsymbol{J})$ of $\boldsymbol{J}$,

$$\{F, G\} = \boldsymbol{J} \cdot \left(\frac{\partial F}{\partial \boldsymbol{J}} \times \frac{\partial G}{\partial \boldsymbol{J}} \right), \tag{2.124}$$

with $\partial F/\partial \boldsymbol{J}$ a convenient shorthand for the vector $\partial F/\partial J_i$ $(i = 1, 2, 3)$. This is, of course, strictly a Poisson manifold and cannot be given a symplectic structure due to the odd dimensionality of angular momentum space, $\dim \Lambda = 3$. This Poisson structure is the unique one such that π is a Poisson map, a map that respects the Poisson brackets of the two Poisson manifolds $\mathcal{S}$ and Λ: $\{F, G\} \circ \pi = \{F \circ \pi, G \circ \pi\}$, see [57]. This construction is the simplest form of reduction possible. The utility of the construction really proceeds in reverse. A calculation with angular momenta can be lifted into the Schwinger phase space $\mathcal{S}$ and often proceeds more simply using the oscillator formalism.

There is a second reduction of the oscillator phase space $\mathcal{S}$ that relates to symplectic reduction and is useful below. A number of perspectives can be taken on this second reduction. Beginning with the space Λ, a general result due to Weinstein [79] states that every finite dimensional Poisson manifold is a disjoint union of symplectic leaves. If there is a Casimir of the bracket, a function that Poisson commutes with everything else, then the Casimir is constant on the symplectic leaves. On those leaves with dimension greater than zero, the Poisson bivector is non-degenerate and can be used to construct the symplectic form. The Hamiltonian I is a Casimir for the bracket on Λ, due to $\{J_i, I\} = 0$ and so the leaves of Λ are two-spheres of constant radius $I = |\boldsymbol{J}|$, since $I^2 = \boldsymbol{J}^2$.

The approach just described proceeds first to Λ and then fixes the energy and arrives at S^2. The symplectic reduction proceeds in a different order from $\mathcal{S}$ to S^2. The reduction begins by considering the level set S^3 determined by a definite value of the oscillator Hamiltonian I, viewed in this context as the momentum map. The orbits generated by I are topological circles, these are the orbits of the isotropy subgroup which in this case is the whole group, and moding out by them is precisely the Hopf map, $\pi_H : S^3 \to S^2$. In coordinates the Hopf map can be described by composing 1/2 the ratio map, which takes $(z_1, z_2) \mapsto z_1/z_2$, with the inverse stereographic projection and this yields

$$\pi_H(z_1, z_2) = \frac{1}{2}(z_1\bar{z}_2 + \bar{z}_1 z_2, i(\bar{z}_1 z_2 - z_1\bar{z}_2), |z_1|^2 - |z_2|^2), \tag{2.125}$$

exactly the Schwinger map of equation (2.122), here, of course, the factor 1/2 is purely to match conventions. Calculating the symplectic form that arises under this reduction yields

$$\omega = Jd\Omega = J\sin(\theta)d\theta d\phi, \tag{2.126}$$

with $J = |\boldsymbol{J}|$ the radius of the sphere and Ω the usual solid angle on the sphere with coordinates (θ, ϕ). This symplectic form is key to the Bohr-Sommerfeld quantization of the volume operator performed in Chapter 4.

We end this chapter as it began, with a final geometrical observation. The Lagrangian manifolds of the 2D oscillator were the tori of constant energy E and energy difference J_3, pictured under stereographic projection in Figure 2.2, and because the Hopf map preserves J_3 these tori project onto circles of constant latitude on the reduced space sphere S^2. This gives a pleasing geometrical picture: as we sweep through the toric leaves of S^3 we build up the reduced space sphere S^2 one circle of latitude at a time, the two poles arising from the degenerate tori with $\theta = 0$ and $\theta = \pi$.

Chapter 3

Spin Networks

3.1 Introduction: simplifying tensor manipulations

Spin networks play a prominent role in the loop quantization of gravity. These objects, which originally arose in the study of the group theory of rotations, provide a convenient basis for the states of the Hilbert space of the theory. In the last section of this chapter (section 3.5) we sketch this relationship in more detail. Spin networks originated from an entire graphical calculus that represents tensors by simple line drawings. The graphical representation is not a panacea, as it does not contain new information not available in the usual coordinate and index notation. However, the cost of learning this new language is quickly repaid by its ability to summarize and simplify the results of tensor manipulations. The graphical calculus also brings closer together the analytical tools of tensors and the results of knot theory, leading to valuable cross-fertilizations.

The central objects of study in this thesis, the spin networks, represent combinations of tensors that are invariant under the action of a group; we will be primarily interested in the double cover of the group of rotations $SU(2)$. Formally, spin networks are defined as graphical representations of intertwiners and intertwiners are linear maps between two vector spaces that commute with a group action on each of these spaces. However, we feel that this formal definition lacks enough context to be useful on a first encounter and doesn't fit into the experience of physicists who haven't worked with spin networks extensively. To bridge this gap section 3.3 roots spin networks in the familiar context of the tensor calculus. This section also explains why the graphical calculus turns out to have so few fundamental elements: the graphical tool box will consist primarily of just two tools, the completely antisymmetric tensors, denoted by index decorated ϵs, and the Kronecker delta δ^a_b.

Our approach to explaining this economy, which at first can seem unreasonable, is to introduce invariant theory. Classically (19th and early 20th centuries) invariant theory was a subject that drew a great deal of attention and was worked on by Cayley, Clebsch and Gordan, Hilbert and many others. Emmy Noether wrote her

dissertation under Paul Gordan on the invariants of homogeneous polynomials in three variables, constructing literally hundreds of invariants by hand. While she subsequently disparaged her dissertation as a "jungle of equations", there is little doubt that this practical experience had an important impact on her discovery, just a few years after her thesis work, of the connection between symmetries and invariants. Nonetheless, the "jungle of equations" sentiment surrounded the theory of invariants for many years. In recent times invariant theory has been rediscovered as a fascinating area of research and it has been recognized that a forked approach applying both abstraction and modern symbolic computation may be sufficient to partially tame the jungle. Several excellent books reintroducing a variety of aspects of the theory have appeared [80, 81, 82].

The second to last section of the chapter, section 3.4, narrows focus and develops our notation and conventions for $SU(2)$ spin networks. The chapter concludes with a discussion of how $SU(2)$ spin networks appear in loop gravity.

3.2 Graphical representation of tensors

To begin we introduce the graphical calculus for $\mathbb{R}^3$. The graphical representation of vectors and tensors treats them as machines whose technical specs are the number and arrangement of indices. This is, in fact, precisely how you hear physicists speak of tensors colloquially and at the blackboard:"the tensor $T_{ab}{}^{c}$ eats two vectors and returns a vector". In the graphical calculus a tensor is represented by a decorated node of a graph and each of its indices is a link coming out of the node. The graphical representations of v^a and $T_{ab}{}^{c}$ are

$$v^a = \quad \overset{a}{\boxed{v}} \qquad T_{ab}{}^{c} = \quad \overset{c}{\underset{a \quad b}{\boxed{T}}} \ . \tag{3.1}$$

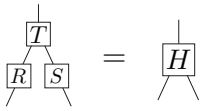
If a particular tensor is used frequently in calculation a more concise representation of the node is preferred over the somewhat bulky $\boxed{T}$. For example we have,

$$\delta^a_b = \quad \overset{a}{\underset{b}{\diagdown}} \qquad \text{and} \qquad \epsilon_{abc} = \quad \overset{a}{\underset{b \quad c}{\curlyvee}} \ , \tag{3.2}$$

for the Kronecker delta and the completely antisymmetric tensor respectively. Typically links exiting the top of a node will be contravariant and those exiting the bottom covariant. This rule is not hard and fast though as is clear from the ϵ_{abc} example above. There is little room for confusion in actual calculations. Index contraction is represented by the joining of two links and as with the Einstein summation convention summation is implied, for example

$$T_{ab}{}^{c} u^a v^b = \quad \overset{c}{\underset{\boxed{u} \quad \boxed{v}}{\boxed{T}}} \ . \tag{3.3}$$

Notice that the labels of the joined links have been dropped. This is because these are now dummy indices and the name of the index is freely chosen. In a similar convention, when graphical equations are written often none of the indices are shown explicitly, instead the position of free links on each side indicates which indices should have the same label, for example,

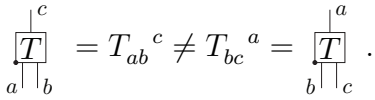


(3.4)

represents $T_{de}{}^{c}R^{d}_{a}S^{e}_{b} = H_{ab}{}^{c}$. The vector dot product $\boldsymbol{v}\cdot\boldsymbol{u}$ is $\boxed{v}-\boxed{u}$ and the trace of the Kronecker delta is

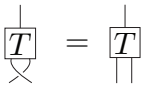
$$= 3. \tag{3.5}$$

Indices are conventionally read in the counter clockwise direction around a tensor node. When it is important to start from a particular index an indication can be made on the tensor node with a dot,



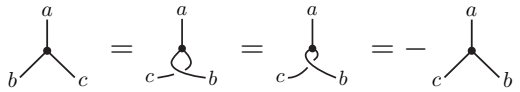
(3.6)

Having laid out the basic correspondences we can examine the manipulation of the graphical representations. The overarching theme is that we can treat these graphs as nodes connected by strings on a table. Any manipulation of the strings correspond to an algebraic manipulation. For example, we can pass two open links over one another,



(3.7)

algebraically this corresponds to having done nothing at all $T_{ab}{}^{c} = T_{ab}{}^{c}$ but notice that you have to read the index ordering off at the socket (where the link meets the node). To see the effect of index transposition the twist of the links has to be carried into the node and, of course, depends on the character of the tensor. So, for example,

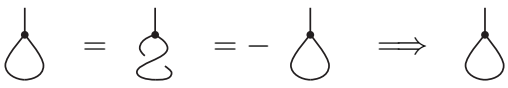


(3.8)

While on the other hand,

$$= 3. \tag{3.9}$$

This demonstrates that the manner in which twists are removed is important in the graphical calculus. As another example we have,

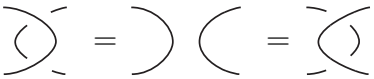


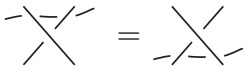
$$= 0. \tag{3.10}$$

This is also clear algebraically, $\epsilon_{abb} = 0$ by the antisymmetry of the symbol. These manipulations exemplify what is meant by a string diagram. The strings of the diagrams are idealized in the sense that they can absorb twists along their length with no effect on the calculations. This is in contrast to a *ribbon* which retains any twists along its length. In this sense ribbons are intrinsically three dimensional, while our string diagrams are planar. There are interesting connections with the q-deformed symbols and ribbon graphs (see e.g. [83]).

We are now in a good position to summarize the allowed diagram moves. The axiomatization presented here is due to Reidemeister [84] and the moves are called Reidemeister moves:

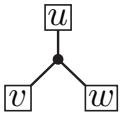
RI:  (3.11)

RII:  (3.12)

RIII:  (3.13)

Two diagrams that are equivalent through a combination of Reidemeister moves are said to be isotopic and give rise to tensor expressions that are equal to one another. Formally this completes the specification of the graphical calculus, however, the great utility of the graphical techniques arises from two additional properties; the easy recognizability of invariants and the graphical reduction rules.

A closed diagram, one with no loose links, is invariant under Euclidean transformations. We've already encountered $\boxed{v} - \boxed{u}$ and another immediate one is,

 $= \boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}).$ (3.14)

Some of these invariants are trivial,

 $= 0,$ (3.15)

but the fact that a quick glance immediately indicates the invariance of an expression is useful. This corresponds to the complete contraction of the indices in a tensor expression, also easily checked but visually more elaborate. Not all invariants are independent, which is intimately connected with the reduction of graphs into smaller pieces, which in turn is algebraically related to the reduction of ϵ tensors.

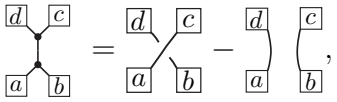
The most useful of these in the case of the Euclidean vector algebra that we are exploring is:

$$\epsilon_{abe}\epsilon^{ecd} = \delta^c_a\delta^d_b - \delta^d_a\delta^c_b, \tag{3.16}$$

graphically this induces the diagram reduction,

(3.17)

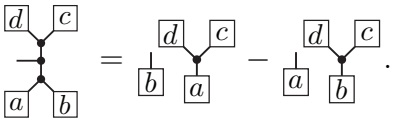
This leads directly to,



(3.18)

$$(\boldsymbol{a}\times\boldsymbol{b})\cdot(\boldsymbol{c}\times\boldsymbol{d})=(\boldsymbol{a}\cdot\boldsymbol{c})(\boldsymbol{b}\cdot\boldsymbol{d})-(\boldsymbol{a}\cdot\boldsymbol{d})(\boldsymbol{b}\cdot\boldsymbol{c}).$$

This reduction move is local, in the sense that it can be applied anywhere in the interior of a more complex graph, and it reduces the number of nodes in the graph. Another example to illustrate the reduction:



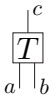
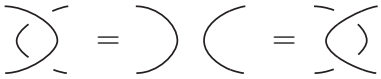
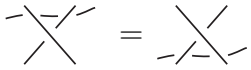
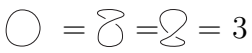
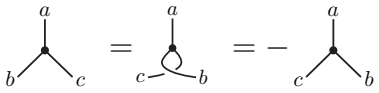
(3.19)

As noted above, any network with all of the strands contracted is an invariant. This corresponds to the familiar claim that any tensor expression with all of the indices contracted is an invariant. This is only a necessary condition and one might wonder whether it is guaranteed that all invariants can be expressed by such diagrams or whether more ingredients are needed in the graphical calculus before we can obtain all invariants. In the next section we take up the investigation of invariants through the tensor calculus and address the sufficiency of the graphical building blocks discussed so far. Table 3.1 summarizes the results of this section and provides a quick reference for doing graphical calculations in $\mathbb{R}^3$.

3.3 Invariant theory

The theory of Clebsch-Gordan coefficients is, at times, presented as a bunch of unmotivated, purely algorithmic rules. In these treatments, the subject has been completely removed from its historical roots. While in some cases this clears the clutter of historical accidents, in this case it makes the theory appear contextless, a tool that is at times applicable but never an inventor's favorite because she doesn't know its insides work. This is unnecessary. In this section we try to recover some of the historical context and show that the coefficients can be cast intuitively in terms of invariants of the more familiar work horse tensors. In fact, there is a beautiful mixture of the algebraic, tensorial and group theoretic tools embodied in Clebsh-Gordan coefficients and in the more symmetrical Wigner symbols that are graphically

Table 3.1: Summary of graphical calculus for $\mathbb{R}^3$

Tensors:	$\delta^a_b =$  $\epsilon_{abc} =$ 
	$v^a =$  $w_a =$  $T_{ab}{}^c =$ 
RI:	
RII:	
RIII:	
Dimension:	
Antisymmetry:	
Reduction:	

represented by spin networks. These tools are combined to extract geometrical insight and become an extremely useful basis for states in loop quantum gravity. This section ends highlighting the role of the group theory and leads into the next section where the focus is narrowed to $SU(2)$ spin networks.

A second purpose of this section will be to explain the economy of the graphical calculus. Here also, a first encounter can leave one uncomfortable. In the case of the graphical calculus, the first impression is often that the rules introduced for the graphical manipulations aren't complex enough to capture the full power of the tensorial machinery to construct invariants. So this section begins with a detailed discussion of how to construct invariants out of tensors and uncovers the entire range of possibilities. We will find that the Kronecker delta, δ, and the completely antisymmetric Levi-Civita symbol, ϵ, play a fundamental role in the construction of invariants. In fact, these are the only tensors that you need in order to construct invariants out of other tensors and this explains the simplicity of the graphical calculus: you only need to introduce graphical representations for δ and ϵ, beyond that the other tensors used to construct invariants are those that arise in the physical application being considered. Connections between invariant theory and tensors were appreciated and developed by Cramlet [85], Littlewood [86], and given a very thorough treatment in the excellent book by Gurevich [87], from which our exposition borrows amply. We begin by introducing more of our conventions with regard to tensors.

3.3.1 Tensorial invariants

In the physics literature tensors are often introduced as multi-index objects with particular transformation properties:

$$T_{\mu\nu}{}^{\rho} \to \tilde{T}_{\mu\nu}{}^{\rho} = M_{\mu}^{\lambda} M_{\nu}^{\sigma} N_{\tau}^{\rho} T_{\lambda\sigma}{}^{\tau}. \tag{3.20}$$

Here we have a tensor of type $(2,1)$, transforming covariantly with respect to its μ- and ν-indices and contravariantly with respect to the ρ-index. More abstractly, the tensor can be viewed as a multilinear map,

$$T : V \otimes V \otimes V^* \to \mathbb{C}, \tag{3.21}$$

and the indexed description is seen as the components of this map after a choice of basis, $\mathcal{B} = \{e_\mu\}_{\mu=1,\ldots,n}$ ($n = \dim V$), has been made in V (with the dual basis, $\mathcal{B}^* = \{e^\nu\}_{\nu=1,\ldots,n}$, being chosen in V^*),

$$T = T_{\mu\nu}{}^{\rho} e^{\mu} \otimes e^{\nu} \otimes e_{\rho}. \tag{3.22}$$

Note that T's definition as the map (3.21) means that T itself lives in the dual space to its domain $T \in V^* \otimes V^* \otimes V$. The matrices M and N are defined as the arrays of numbers that express the basis $\tilde{\mathcal{B}} = \{\tilde{e}_\mu\}_{\mu=1,\ldots,n}$ in terms of the basis $\mathcal{B}$ (and similarly

for duals), that is $\tilde{e}_\mu = M^\rho_\mu e_\rho$ and $\tilde{e}^\nu = N^\nu_\rho e^\rho$.[1] As matrices M and N are inverses as is clear from

$$\tilde{e}^\nu(\tilde{e}_\mu) = \delta^\nu_\mu = N^\nu_\sigma M^\rho_\mu e^\sigma(e_\rho) = N^\nu_\rho M^\rho_\mu. \tag{3.23}$$

Expressing T in the tilde basis reproduces (3.20),

$$\tilde{T}_{\mu\nu}{}^\rho = T(\tilde{e}_\mu, \tilde{e}_\nu, \tilde{e}^\rho) = T(M^\lambda_\mu e_\lambda, M^\sigma_\nu e_\sigma, N^\rho_\tau e^\tau) = M^\lambda_\mu M^\sigma_\nu N^\rho_\tau T_{\lambda\sigma}{}^\tau. \tag{3.24}$$

In this section we are interested in the most general invertible transformations, so the matrix M will be drawn from the general linear group $M \in GL_n(\mathbb{C})$. This broad context will lead to consideration of a slight generalization of a tensor known as a tensor density. Tensor densities are sensitive to overall changes in scale and transform as tensors except for a power of the determinant $\Delta \equiv \det M$,

$$\tilde{E}_{\mu\nu}{}^\rho = \Delta^w M^\lambda_\mu M^\sigma_\nu N^\rho_\tau E_{\lambda\sigma}{}^\tau, \tag{3.25}$$

here the power w is called the *weight* of the tensor density and we will assume that the weight is always an integer (negative or positive), $w \in \mathbb{Z}$. The tensor density of greatest interest in what follows, ϵ, is modeled on the volume form $e = e^1 \wedge \cdots \wedge e^n$, where the $\wedge$ is the antisymmetric wedge product. The components of e

$$e(e_{\mu_1}, \cdots, e_{\mu_n}) = e_{\mu_1 \cdots \mu_n} \tag{3.26}$$

are related to the single essential component $e_{1\cdots n}$ (again $e^{\nu_1}(e_{\mu_1}) = \delta^{\nu_1}_{\mu_1}$). This is because if $\mu_i = \mu_j$ $(i \neq j)$ for any pair of indices μ_i and μ_j then the component with those indices vanishes due to the antisymmetry of the wedge product. All other index sets are permutations of $\{1, \ldots, n\}$. Index sets that are even or odd permutations of this set give rise to components that can be permuted to obtain $\pm e_{1\cdots n}$ respectively. Under the change of basis from $\mathcal{B}$ to $\tilde{\mathcal{B}}$ we have

$$\begin{aligned} \tilde{e}_{1\cdots n} &= e(\tilde{e}_1, \cdots, \tilde{e}_n) = M^{\mu_1}_1 \cdots M^{\mu_n}_n e_{\mu_1 \cdots \mu_n} \\ &= \left(\sum_{\pi \in \mathcal{I}} \mathrm{sgn}(\pi) M^{\pi(\mu_1}_1 \cdots M^{\mu_n)}_n \right) e_{1\cdots n} = \Delta\ e_{1\cdots n}, \end{aligned} \tag{3.27}$$

here we have introduced a bit of notation: π is a permutation, $\mathcal{I}$ is the set of all permutations of $\{1, \cdots, n\}$, $\mathrm{sgn}(\pi)$ is ± 1 if π is an even, respectively odd permutation

[1] We adopt the passive conventions here because we have general relativity in mind. The general covariance of relativity is, perhaps, most easily thought of as a freedom to choose whatever coordinates you like to describe the physics at hand. To our sensibilities the condition of invariance, introduced below, is more naturally expressed in terms of active transformations. The cylindrical symmetry of a vase is clearest when you pick it up, rotate it and put it back down. But, of course, it is completely equivalent to walk around the vase. Anyway, a choice must be made and we stick to the passive transformations throughout this treatment.

and $\pi(\mu_1 \cdots \mu_n)$ means $\mu_{\pi(1)} \cdots \mu_{\pi(n)}$. The last equality follows from the definition of the determinant

$$\Delta \equiv \det M \equiv \sum_{\pi \in \mathcal{I}} \mathrm{sgn}(\pi) M_1^{\pi(\mu_1} \cdots M_n^{\mu_n)}. \tag{3.28}$$

So this single component transforms as a scalar density of weight $+1$.

Now, define ϵ as the tensor density of weight -1, with essential component $\epsilon_{1\cdots n} = 1$ and

$$\epsilon_{\mu_1 \cdots \mu_n} = \begin{cases} +1 & \text{if } \mu_1 \cdots \mu_n \text{ is an even permutation of } 1 \cdots n, \\ 0 & \text{if } \mu_i = \mu_j \text{ for } i \neq j, \\ -1 & \text{if } \mu_1 \cdots \mu_n \text{ is an odd permutation of } 1 \cdots n. \end{cases} \tag{3.29}$$

Because ϵ is defined to *transform* as a tensor density of weight -1, its components are invariant:

$$\begin{aligned} \tilde{\epsilon}_{1\cdots n} &= \Delta^{-1}\ M_1^{\mu_1} \cdots M_n^{\mu_n} \epsilon_{\mu_1 \cdots \mu_n} \\ &= \Delta^{-1} \Delta\ \epsilon_{1\cdots n} = \epsilon_{1\cdots n}, \end{aligned} \tag{3.30}$$

and so the equation (3.29) holds in any coordinate system. We refer to this tensor density as the covariant Levi-Civita symbol or more loosely as an ϵ-tensor. In a parallel manner we introduce the tensor density $\epsilon^{\mu_1 \cdots \mu_n}$ with weight $+1$ and invariant components. These definitions are admittedly contorted, however they allow the introduction of these tensor densities with no reference to a metric on V or V^*.

As emphasized by, for example, Penrose and Wald it is also useful to have an intermediate notation $T_{ab}{}^c$ where now the indices are abstract. They tell you the order (total number of indices) and type of a tensor but are not referred to any particular choice of basis. This abstract index structure is mirrored in the graphical calculus where the number of strands coming out of a node and their position tells you the same information. The abstract index notation was implicitly used throughout the last section. In a slight abuse of notation we will also write M_a^b and N_c^d. The two collections of numbers M and N are defined with respect to the bases that they connect and so it seems odd to speak of them without reference to basis. Nonetheless, no serious inconsistencies are introduced in this manner, they are tensors and the important property (3.23) doesn't depend on what bases they connect. Subsequently we will freely transition between the various ways of representing tensors (for more discussion of tensors and some of their subtleties see the excellent new introduction to tensors [88]).

Before commencing discussion of invariant tensors we review some properties of invariants in general. The invariants considered here will be functions of the components of tensors. The arguments of an invariant I will be denoted with lower case letter, e.g. the components of $t_{a_1 \cdots a_n}{}^{b_1 \cdots b_n}$. Also, because writing $I(t_{1\cdots 1}{}^{1\cdots 1}, \ldots, t_{n\cdots n}{}^{n\cdots n})$ is awkward, instead we write $I(t)$. Of course, a general invariant I can depend on

the components of an arbitrary number of tensors or tensor densities. An absolute invariant $I(t)$ is defined by the condition

$$I(\tilde{t}) = I(t), \tag{3.31}$$

where t has been subjected to the linear transformation

$$\tilde{t}_{a_1\cdots a_r}{}^{b_1\cdots b_s} = M^{c_1}_{a_1}\cdots M^{c_r}_{a_r} N^{b_1}_{d_1}\cdots N^{b_s}_{d_s} t_{c_1\cdots c_r}{}^{d_1\cdots c_s}. \tag{3.32}$$

This is in contrast to a relative invariant, defined by

$$I(\tilde{t}) = \Delta^w I(t), \tag{3.33}$$

where once again $\Delta \equiv \det M$ and w is called the weight of the invariant. Note that relative invariants become absolutely invariant when attention is restricted to subgroups of $GL_n(\mathbb{C})$ with unit determinant, the *special* subgroups. The latter definition is less strict than it appears. Let I be a rational function (ratio of polynomials) of the components of t and suppose it satisfies

$$I(\tilde{t}) = f(M^b_a) I(t), \tag{3.34}$$

with f an arbitrary function. Then it can be shown that $f(M) = \Delta^w$ for an integer w (recall that we have assumed that all tensor densities have integer weights). The introduction of rational functions here is not arbitrary. The most general invariant considered here will be an algebraic invariant I, that is, an invariant satisfying a polynomial equation

$$[I(t)]^k + C_1(t)\,[I(t)]^{k-1} + \cdots + C_k(t) = 0, \tag{3.35}$$

with $C_1, \ldots, C_k$ rational functions of the tensor components. We assume this polynomial is irreducible so that k is the least such power and the coefficients $C_1, \ldots, C_n$ are unique. It follows that the coefficients $C_1, \ldots, C_k$ are rational invariants and that the weight of I is a rational number. Thus all the invariants investigated here are determined by rational invariants with integer weights. Finally any rational invariant can be expressed as a sum of homogeneous rational invariants, where the homogeneity is in the power of the components of t. all of these results have straight forward demonstrations (see section 15 of [87]) but they are lengthy enough to be a distraction here. As emphasized by Olver ([80]) there are many connections between the algebraic invariants considered here and differential invariants such as the Jacobian and Hessian. These definitions and results establish the foundation for the central results of this section involving invariant tensors.

An (absolutely) invariant tensor is a tensor whose components are all invariant under linear transformation. Such a tensor necessarily has type (r, r). This can

be seen by considering a particular linear transformation that is the scaling of the identity $M^b_a = \lambda\delta^b_a$ with $\lambda > 0$. The invariance condition

$$\tilde{T}_{a_1\cdots a_r}{}^{b_1\cdots b_s} = M^{c_1}_{a_1}\cdots M^{c_r}_{a_r}N^{b_1}_{d_1}\cdots N^{b_s}_{d_s}T_{c_1\cdots c_r}{}^{d_1\cdots d_s} = T_{c_1\cdots c_r}{}^{d_1\cdots d_s} \tag{3.36}$$

cannot be satisfied unless there are an equal number of Ns (with $N^b_a = 1/\lambda\delta^b_a$) transforming T. Next consider the invariance condition for a general type (r, r) tensor and a general transformation M. Due to $N = M^{-1}$ (see (3.23)) it can be written

$$M^{c_1}_{a_1}\cdots M^{c_r}_{a_r}T_{c_1\cdots c_r}{}^{d_1\cdots d_r} = M^{d_1}_{b_1}\cdots M^{d_r}_{b_r}T_{a_1\cdots a_r}{}^{b_1\cdots b_r} \tag{3.37}$$

To illustrate the general procedure used below, specialize to the type $(2, 2)$ case:

$$M^{c_1}_{a_1}M^{c_2}_{a_2}T_{c_1c_2}{}^{d_1d_2} = M^{d_1}_{b_1}M^{d_2}_{b_2}T_{a_1a_2}{}^{b_1b_2}. \tag{3.38}$$

Differentiating this identity with respect to $M^{q_1}_{p_1}$ and the result with respect to $M^{q_2}_{p_2}$ yields,[2]

$$\delta^{p_1}_{a_1}\delta^{p_2}_{a_2}T_{q_1q_2}{}^{d_1d_2} + \delta^{p_2}_{a_1}\delta^{p_1}_{a_2}T_{q_2q_1}{}^{d_1d_2} = \delta^{d_1}_{q_1}\delta^{d_2}_{q_2}T_{a_1a_2}{}^{p_1p_2} + \delta^{d_1}_{q_2}\delta^{d_2}_{q_1}T_{a_1a_2}{}^{p_2p_1}. \tag{3.39}$$

Setting $a_1 = p_1 = 1$ and $a_2 = p_2 = 2$ yields,

$$\begin{aligned} T_{q_1q_2}{}^{d_1d_2} &= T^{12}_{12}\delta^{d_1}_{q_1}\delta^{d_2}_{q_2} + T^{21}_{12}\delta^{d_1}_{q_2}\delta^{d_2}_{q_1} \\ &= \sum_{\pi\in\mathcal{I}} C_\pi\,\delta^{\pi(d_1}_{q_1}\delta^{d_2)}_{q_2} \end{aligned} \tag{3.40}$$

in the second equality the notation is the same as that introduced below (3.27) and C_π is shorthand for the constant associated with that permutation, here we have $C_{12} = T^{12}_{12}$ and $C_{21} = T^{21}_{12}$. This shows that the only absolutely invariant tensors of type $(2, 2)$ consist of sums of tensors $\delta^{b_1}_{a_1}\delta^{b_2}_{a_2}$ with arbitrary constant coefficients. This extends straight forwardly to type (r, r) tensors, differentiating (3.37) with respect to $M^{q_1}_{p_1}, \ldots, M^{q_r}_{p_r}$ sequentially yields

$$\sum_{\pi\in\mathcal{I}}\delta^{\pi(p_1}_{a_1}\cdots\delta^{p_r)}_{a_r}T_{\pi(q_1\cdots q_r)}{}^{d_1\cdots d_r} = \sum_{\pi\in\mathcal{I}}\delta^{d_1}_{\pi(q_1}\cdots\delta^{d_r}_{q_r)}T_{a_1\cdots a_r)}{}^{\pi(d_1\cdots d_r)} \tag{3.41}$$

and setting $a_1 = p_1 = 1, \ldots, a_r = p_r = r$ we find

$$T_{q_1\cdots q_r}{}^{d_1\cdots d_r} = \sum_{\pi\in\mathcal{I}} C_\pi\delta^{d_1}_{\pi(q_1}\cdots\delta^{d_r}_{q_r)}. \tag{3.42}$$

We have implicitly been assuming that $n = \dim V \geq r$, the result still holds for $r > n$ although some care is needed and the constants C_π are no longer unique. The non-uniqueness of the C_π is because for $r > n$ the product of δs, $\delta^{d_1}_{q_1}\cdots\delta^{d_r}_{q_r}$,

[2]The differentiation trick is due to Appleby et al [89].

necessarily has a repeated term and so any multiple of $\delta^{d_1}_{[q_1}\cdots\delta^{d_r}_{q_r]}$ ($[\cdots]$ denoting antisymmetrization) can be added to the right hand side of (3.42) changing the C_π but leaving $T_{q_1\cdots q_r}{}^{d_1\cdots d_r}$ unaffected. In conclusion, products of δs are the only absolute invariant tensors.

This classification extends to relative tensors in a simple manner. Assume that $R_{a_1\cdots a_r}{}^{b_1\cdots b_s}$ is a relative tensor of weight w. Transforming with $M^d_c = \lambda\delta^d_c$, $\Delta = \lambda^n$ and

$$\tilde{R}_{a_1\cdots a_r}{}^{b_1\cdots b_s} = \Delta^w\lambda^{r-s}R_{a_1\cdots a_r}{}^{b_1\cdots b_s} = \lambda^{nw+r-s}R_{a_1\cdots a_r}{}^{b_1\cdots b_s}. \tag{3.43}$$

and requiring invariance imposes

$$r + nw = s. \tag{3.44}$$

Suppose that a relative invariant tensor R of positive weight w is given, then an absolute invariant can be constructed from it by multiplying by w Levi-Civita symbols,

$$R_{a_1\cdots a_r}{}^{b_1\cdots b_s}\epsilon_{a_{r+1}\cdots a_{r+n}}\cdots\epsilon_{a_{r+1+nw-n}\cdots a_{r+wn}}. \tag{3.45}$$

Because the ϵ are weight -1 this combination transforms as a tensor. Also due to the assumption on R and the invariance of the ϵs this is an absolute invariant and hence expressible as the sum

$$R_{a_1\cdots a_r}{}^{b_1\cdots b_s}\epsilon_{a_{r+1}\cdots a_{r+n}}\cdots\epsilon_{a_{s-n+1}\cdots a_s} = \sum_{\pi\in\mathcal{I}_s}\widehat{C}_\pi\delta^{\pi(b_1}_{a_1}\cdots\delta^{b_s)}_{a_s}, \tag{3.46}$$

we've used (3.44) to simplify the indices. Contracting both sides of this equation with $\epsilon^{a_{r+1}\cdots a_{r+n}}\cdots\epsilon^{a_s-n+1\cdots a_s}$ and dividing through by $(n!)^w$ yields

$$\begin{aligned} R_{a_1\cdots a_r}{}^{b_1\cdots b_s} &= \sum_{\pi\in\mathcal{I}_s} C_\pi\delta^{\pi(b_1}_{a_1}\cdots\delta^{b_s)}_{a_s}\epsilon^{a_{r+1}\cdots a_{r+n}}\cdots\epsilon^{a_s-n+1\cdots a_s} \\ &= \sum_{\pi\in\mathcal{I}_s} C_\pi\delta^{\pi(b_1}_{a_1}\cdots\delta^{b_r}_{a_r}\epsilon^{b_{r+1}\cdots b_{r+n}}\cdots\epsilon^{b_{s-n+1}\cdots b_s)} \end{aligned} \tag{3.47}$$

with $C_\pi \equiv \widehat{C}_\pi/(n!)^w$. Again a parallel argument in the negative weight case ($r > s$) yields

$$R_{a_1\cdots a_r}{}^{b_1\cdots b_s} = \sum_{\pi\in\mathcal{I}_s} C_\pi\delta^{b_1}_{\pi(a_1}\cdots\delta^{b_s}_{a_s}\epsilon_{a_{s+1}\cdots a_{s+n}}\cdots\epsilon_{a_{r-n+1}\cdots a_r)} \tag{3.48}$$

with $r = s + n|w|$. All invariant tensor densities can be expressed as a linear combination of Kronecker deltas and Levi-Civita symbols.

With the classification of invariant tensors settled we turn to our main argument: All algebraic (absolute and relative) invariants are formed by contraction with invariant tensors.

As argued above, consideration need only be given to homogeneous rational invariants. Suppose such an invariant I is to be constructed out of the components of the

tensors t, u and v (an arbitrary number of tensors could be considered). Each term of the invariant consists of a numerical coefficient times a product of the components of t, u and v. Decompose I as,

$$I = C_{a_1\cdots a_r}{}^{b_1\cdots b_s} K_{b_1\cdots b_s}{}^{a_1\cdots a_r}, \tag{3.49}$$

where K is the collection of the components of t, u and v making up the terms of I and C is an array containing the coefficients multiplying these components to create the invariant. Without loss of generality the coefficient array C can be taken to have the same symmetries as the component tensor K. If I is a relative invariant of weight w the condition of invariance is

$$C_{a_1\cdots a_r}{}^{b_1\cdots b_s} \tilde{K}_{b_1\cdots b_s}{}^{a_1\cdots a_r} = \Delta^w C_{a_1\cdots a_r}{}^{b_1\cdots b_s} K_{b_1\cdots b_s}{}^{a_1\cdots a_r}. \tag{3.50}$$

Assuming the component tensor K is of weight w_k then this becomes

$$\Delta^{w_k - w} C_{a_1\cdots a_r}{}^{b_1\cdots b_s} M_{b_1}^{c_1}\cdots M_{b_s}^{c_s} N_{d_1}^{a_1}\cdots N_{d_r}^{a_r} K_{c_1\cdots c_s}{}^{d_1\cdots d_r} = C_{a_1\cdots a_r}{}^{b_1\cdots b_s} K_{b_1\cdots b_s}{}^{a_1\cdots a_r}. \tag{3.51}$$

Defining a (relative) tensor of weight $w_c = w - w_k$ with components given by the array C, the factors on the left hand side of (3.51) multiplying K can be seen as a transformed version $\bar{C}$ of C. This interpretation requires the inverse transformation of the one usually considered, that is the one taking e_μ to $\bar{e}_\mu = N_\mu^\nu e_\nu$, with $\bar{e}_\mu$ a member of the basis $\bar{\mathcal{B}} = \{\bar{e}_\mu\}_{\mu=1,\ldots,n}$. Because $\det N = \Delta^{-1}$, the factor $\Delta^{w_k - w}$ can be seen as the weight factor of this inverse transformation. Expressed in terms of this tensor density the invariance condition is

$$\left(\tilde{C}_{a_1\cdots a_r}{}^{b_1\cdots b_s} - C_{a_1\cdots a_r}{}^{b_1\cdots b_s}\right) K_{b_1\cdots b_s}{}^{a_1\cdots a_r} = 0. \tag{3.52}$$

This is a polynomial in the components of the tensors t, u and v and for arbitrary such components the coefficients must vanish. But the vanishing of these coefficients is the invariance condition for C,

$$\tilde{C}_{a_1\cdots a_r}{}^{b_1\cdots b_s} = C_{a_1\cdots a_r}{}^{b_1\cdots b_s}. \tag{3.53}$$

Then, according to the results just proved on invariant tensors C is a linear combination of products of δs and ϵs. It follows that all homogeneous invariants $I(t, u, v\ldots)$ are constructed by total contraction of Kronecker δs and Levi-Civita ϵs with the tensors of interest $t, u, v, \ldots$ and because of the relationship between homogeneous rational invariants and algebraic invariants we see that all algebraic invariants are ultimately determined by such contractions.

This, somewhat lengthy, detour into invariant theory provides several benefits. The graphical calculus, which begins with the introduction of the diagrams for δ and ϵ tensors, can now be confidently applied to obtain all invariants. The arguments of

this section have in them the seeds of other important theorems such as the Wigner-Eckart theorem, which also have very nice graphical formulations [83, 90, 91]. This approach also alleviates a difficulty of the graphical formulation, namely that it can be hard to be sure that all graphs of interest have been enumerated. When this difficulty arises one can turn to the algebraic formulation. The explosion of indices in this section also indicates the advantage of the index free graphical notation. So far we have explored some generalities of invariant theory, motivating the economy of the graphical representation. Now, we turn to some concrete examples. The examples will help to clarify the previous discussion and they will also crop up in surprising ways in the discussion of particular spin networks given in later chapters.

3.3.2 Examples: Euclidean invariants & discriminants

Already in section 3.2 we have encountered the basic invariants of Euclidean $\mathbb{R}^3$, $\boldsymbol{u}\cdot\boldsymbol{v}$ and $\boldsymbol{u}\cdot(\boldsymbol{v}\times\boldsymbol{w})$. To what extent do these exhaust the possible invariants?[3] Let us focus for a moment on all invariants constructed from the two arbitrary vectors $\boldsymbol{u}$ and $\boldsymbol{v}$. By rotating our coordinate system we can arrange for $\boldsymbol{u}$ to lie along the x-axis and $\boldsymbol{v}$ to lie in the xy-plane. With this choice their coordinate expressions become $(u_x, 0, 0)$ and $(v_x, v_y, 0)$ respectively. These coordinate expressions can be written in terms of the dot products as $u_x = \sqrt{\boldsymbol{u}\cdot\boldsymbol{u}}$, $v_x = \boldsymbol{u}\cdot\boldsymbol{v}/\sqrt{\boldsymbol{u}\cdot\boldsymbol{u}}$ and $v_y = \sqrt{\boldsymbol{v}\cdot\boldsymbol{v} - (\boldsymbol{u}\cdot\boldsymbol{v})^2/(\boldsymbol{u}\cdot\boldsymbol{u})}$. Because we can always adopt such a coordinate system all invariants, I, built out of $\boldsymbol{u}$ and $\boldsymbol{v}$ are expressible in terms of u_x, v_x and v_y, $I = I(u_x, v_x, v_y)$. Under a parity transformation of each axis $u_x \mapsto -u_x$, $v_x \mapsto -v_x$ and $v_y \mapsto -v_y$ and for I to be invariant under both these transformations it must consist of an even power of v_y and either even powers of u_x and v_x or of powers of their product $u_x v_x$. These powers are all expressible in terms of rational functions of the dot products, i.e. all of the square roots can be eliminated from the invariant.

More generally, for n vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ all invariants are functions of the n^2 invariants

$$\begin{pmatrix} \boldsymbol{v}_1^2 & \boldsymbol{v}_1\cdot\boldsymbol{v}_2 & \cdots & \vdots \\ \boldsymbol{v}_2\cdot\boldsymbol{v}_1 & \boldsymbol{v}_2^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots & \boldsymbol{v}_n^2 \end{pmatrix}. \tag{3.54}$$

This matrix of dot products is called a Gram matrix and is very useful in the study of the $3nj$-symbols, more on Gram matrices in Chapter 5. The general result would be difficult to derive using the coordinate based arguments of the last paragraph. Weyl uses a tool from the old formalism of invariant theory [92]. Instead, here we will combine the graphical tools of section 3.2 with the invariant theory arguments of the last section. Above we showed that all invariants can be written as closed networks

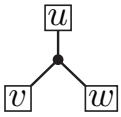
[3] The simple argument here is taken from Weyl's book [92].

contracted with ϵ-tensors. Graphs containing more than two ϵ-tensors can be reduced using the basic identity

(3.55)

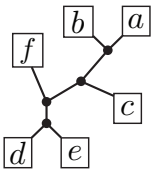
until sums of products of graphs containing only zero or one node (ϵ-tensor) remain. So to show the Gram matrix result we turn to networks containing a single node.

A closed network with one epsilon node is a triple product,



$$= \boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}). \tag{3.56}$$

Geometrically it represents the volume of the parallelepiped spanned by $\boldsymbol{u}, \boldsymbol{v}$ and $\boldsymbol{w}$ or if normalized properly the volume, V, of a tetrahedron with $\boldsymbol{u}, \boldsymbol{v}$ and $\boldsymbol{w}$ as edge vectors, $V = \frac{1}{6}\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})$. To show that a single node can also be expressed in terms of dot products consider the following closed network consisting of the six vectors $\boldsymbol{a}, \ldots, \boldsymbol{f}$,

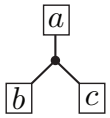


(3.57)

The following calculation is best done by hand: apply reduction on the lower two nodes and then apply it again on the upper two nodes and you will find that the graph reduces to the product of two trivalent networks plus several dot products in the first case and just a sum of dot products in the second case . Solving these two equations for the product of trivalent nodes you find that this product is expressed completely in terms of dot products. In this case the result is easily derived algebraically, despite being more fun graphically; let R and S be the matrices whose columns are the vectors $\boldsymbol{a}, \boldsymbol{b}$ and $\boldsymbol{c}$ and $\boldsymbol{d}, \boldsymbol{e}$ and $\boldsymbol{f}$ respectively then

$$\det R \det S = \det R^T \det S = \det R^T S = \det \begin{pmatrix} \boldsymbol{a}\cdot\boldsymbol{d} & \boldsymbol{a}\cdot\boldsymbol{e} & \boldsymbol{a}\cdot\boldsymbol{f} \\ \boldsymbol{b}\cdot\boldsymbol{d} & \boldsymbol{b}\cdot\boldsymbol{e} & \boldsymbol{b}\cdot\boldsymbol{f} \\ \boldsymbol{c}\cdot\boldsymbol{d} & \boldsymbol{c}\cdot\boldsymbol{e} & \boldsymbol{c}\cdot\boldsymbol{f} \end{pmatrix}, \tag{3.58}$$

however it's nice to see the manner in which the derivation is forced on you by the graphical techniques, all you could do was apply reduction. This is particularly useful when you do not already know the appropriate mathematical structure, such as the determinant, to use to write the invariant. Specializing to $\boldsymbol{a} = \boldsymbol{d}, \boldsymbol{b} = \boldsymbol{e}, \boldsymbol{c} = \boldsymbol{f}$ yields,



$$= \pm\sqrt{\det \begin{pmatrix} \boldsymbol{a}^2 & \boldsymbol{a}\cdot\boldsymbol{b} & \boldsymbol{a}\cdot\boldsymbol{c} \\ \boldsymbol{b}\cdot\boldsymbol{a} & \boldsymbol{b}^2 & \boldsymbol{b}\cdot\boldsymbol{c} \\ \boldsymbol{c}\cdot\boldsymbol{a} & \boldsymbol{c}\cdot\boldsymbol{b} & \boldsymbol{c}^2 \end{pmatrix}}. \tag{3.59}$$

Therefore all invariants constructed out of n vectors in $\mathbb{R}^3$ can be expressed as products and square roots of products of dot products.

A number of questions about this theorem immediately arise: How many distinct invariants are there at a certain order (and number of vectors)? Are there algebraic relationships between the various invariants? Addressing these issues will be the focus of the next section.

There is more to say about the Gram matrix result (3.59). The dot products can be expressed in terms of the magnitudes of the vectors and their differences,

$$\boldsymbol{a} \cdot \boldsymbol{b} = \frac{1}{2}(a^2 + b^2 - (\boldsymbol{a} - \boldsymbol{b})^2). \tag{3.60}$$

Inserting this expression into the Gram matrix result shows that the volume of a tetrahedron can be expressed as a determinant of the squared lengths of the edges of the tetrahedron. There is a very nice way to reorganize the determinant expressed in terms of edge lengths first found by Cayley. Building on Cayley's work Menger formulated an axiomatization of metric geometry completely in terms of lengths, these days this is known as distance geometry and a very thorough introduction is available in [93] and on the more applied side [94]. This way of looking at things forges a nice connection with Klein's Erlangen program in which geometry is viewed as arising from the invariants of a group, in this case the lengths of objects in $\mathbb{R}^3$ under the Euclidean transformations. The Cayley-Menger determinant was used extensively by Ponzano and Regge in their seminal work [11], but we have found it both conceptually and computationally simpler to work with the Gram matrix in the context of spin networks.

Another important and broad ranging application of invariant theory is the subject of discriminants. Given a polynomial,

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0 \tag{3.61}$$

with roots $r_1, \cdots, r_n$, the discriminant of the polynomial is defined by

$$D(P) = a_n^{2n-2} \prod_{i<j} (r_i - r_j)^2. \tag{3.62}$$

From this definition it is clear that the discriminant vanishes whenever two or more of the roots of P coalesce. Because the roots only depend on the coefficients of the polynomial, D can also be seen as a function of these coefficients. Astoundingly expressions for D in terms of the $a_0, \ldots, a_n$ can be found without ever finding the roots $r_1, \ldots, r_n$. There are a variety of ways of doing this and a fun constructive route is to use the graphical calculus. We briefly illustrate with the quadratic case, leaving the more interesting cubic and quartic cases open for exploration, (see [95] for the results).

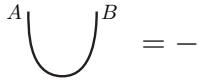
To set things up we convert the quadratic equation $ax^2 + bx + c = 0$ into a homogeneous form,

$$P(x, y) = ax^2 + bxy + cy^2 \tag{3.63}$$

and view the coefficients as a tensor Q_{AB} contracted with $\xi^A = (x, y)$,

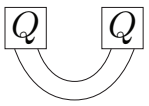
$$P = Q_{AB}\xi^A\xi^B, \qquad Q_{AB} = \begin{pmatrix} a & 1/2b \\ 1/2b & c \end{pmatrix}. \tag{3.64}$$

The idea is to look for invariants constructed out of the tensor Q_{AB}. We represent the two dimensional Levi-Civita ϵ_{AB} by



$$= -\epsilon_{AB} \tag{3.65}$$

(see the next section for an explanation of this representation, including the sign). A little scribbling with string diagrams lands on the simple invariant,



.

This is twice the determinant of Q_{AB} and indeed,

$$\epsilon^{AD}\epsilon^{BC}Q_{AB}Q_{CD} = 2\left(ac - \frac{b^2}{4}\right) = -\frac{1}{2}D_2, \tag{3.66}$$

where D_2 is the discriminant of the quadratic equation. The advantage of the graphical technique is that it is easy to *see* the basic invariants. In the case of the quartic equation you will find that the discriminant is a polynomial in the basic graphical invariants you construct. Algebraic relations between invariants like this are called syzygies and will be touched on below.

Another way of constructing discriminants is to build on a geometrical insight. The two roots of a polynomial coalesce if the first derivative of the polynomial vanishes at the same time as the polynomial. This idea leads into the theory of resultants ([82]) which we will not discuss here except to say that it yields a very simple formula for the discriminant; the $(2n-1) \times (2n-1)$ Sylvester matrix A_S is defined as

$$A_S = \begin{pmatrix} a_n & a_{n-1} & \cdots & a_0 & 0 & \cdots & \cdots & 0 \\ 0 & a_n & a_{n-1} & \cdots & a_0 & 0 & \cdots & 0 \\ \vdots & & & & & & & \vdots \\ 0 & \cdots & 0 & a_n & a_{n-1} & \cdots & \cdots & a_0 \\ na_n & (n-1)a_{n-1} & \cdots & 1a_1 & 0 & \cdots & \cdots & 0 \\ 0 & na_n & (n-1)a_{n-1} & \cdots & 1a_1 & 0 & \cdots & 0 \\ \vdots & & & & & & & \vdots \\ 0 & \cdots & 0 & na_n & (n-1)a_{n-1} & \cdots & \cdots & 1a_1 \end{pmatrix} \tag{3.67}$$

and the discriminant is

$$D = (-1)^{\frac{1}{2}n(n-1)} \frac{1}{a_n} \det A_S. \tag{3.68}$$

Discriminants can be useful in semiclassical theory. If a Lagrangian manifold is described by the vanishing of a polynomial equation then its caustic set can be found as the discriminant of this polynomial because this is precisely where the branches of the manifold coalesce. This is how the caustic set of the $9j$-symbol is found in Chapter 5. In principal this technique can be extended to more complex Lagrangian manifolds (algebraic varieties) although the calculation of resultants of systems of equations quickly becomes infeasible.

3.3.3 Abstract formulation

Our approach to invariant theory has been to get our hand dirty with coordinate and component calculations, after all we have applications in mind. However, an abstract approach is well developed and quite beautiful. We briefly review the results of this approach to situate what we have discussed in the broader context of what is possible (for introductions see [81, 80]).

In section **??** we discussed the fact that a general algebraic invariant is determined by homogeneous rational invariants. We also showed that a rational invariant differs from a polynomial one by a power of the determinant of the linear transformation being considered. For these reasons the abstract formulation focuses on homogeneous polynomials. Let G be an arbitrary group and suppose a linear representation ρ of G on V is given. If $(t_1, \ldots, t_n)$ are coordinates of V (in the basis $\mathcal{B}$) and $I(t_1, \ldots, t_n)$ is a polynomial in these coordinates then if I is unchanged by the transformations $\rho(g)$ for all $g \in G$ then I is an invariant. We have been considering the matrix representations $M^b_a \in GL_n(\mathbb{C})$.

At the turn of the 19th century Hilbert proved two fundamental theorems in invariant theory and as a side effect launched modern abstract algebra. The first of these theorems states that the algebra of invariants of a given representation of G is finitely generated. This means that all invariants ϕ (of this representation of G) can be written as polynomials $\phi(I_1, \ldots, I_n)$ of a finite set of fundamental invariants $\{I_1, \ldots, I_m\}$.

Generically the invariants $\{I_1, \ldots, I_m\}$ are not independent but satisfy d algebraic identities $Z_s(I_1, \ldots, I_m) = 0$, $(s = 1, \ldots, d)$. These identities are called syzygies and Hilbert's second fundamental theorem proves that the set of syzygies is also finite.[4] More precisely there can be secondary syzygies amongst the original set and so on down the line. The complete set of all of these relations is finite.

Hilbert's original proof of the first fundamental theorem was not constructive and was criticized on these grounds. He found another, constructive proof, of the theorem

[4]Syzygy comes from the late Latin syzygia literally meaning yolked together and is also used in astronomy in the sense of conjunction.

a few years later. In recent times the Gröbner basis technology has become a practical tool for calculating systems of invariants, see [81]. Computer algebra packages, such as Mathematica, can be used for these calculations as long as there are not too many variables and the invariants are not too high order. Mathematicians are tackling the surrounding territory anew: they are working on techniques for solving systems of polynomial equations [96]; exploring discriminants, resultants and hyperdeterminants [82]; and making forays into a new non-linear algebra [97]. As this work becomes practicable it will have a huge impact on the physical sciences, where non-linear processes are the rule.

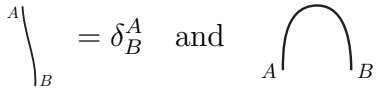
In an obvious but nonetheless interesting manner, restricting attention to a subgroup of the general linear group increases the number of invariants. This has already cropped up several times in our discussion of tensor densities. As soon as we consider the special linear subgroup all relative invariants become absolute invariants with respect to this group. Depending on the group chosen the character of the invariants varies. Weyl's book *The Classical Groups*, [92], is a detailed exploration of the range of possibilities for the symmetric, orthogonal and symplectic groups. The thrust of the next section is to build a complete understanding of the invariants of $SU(2)$. Again the Levi-Civita tensors play a foundational role in building up these invariants but their diagrammatic description will give way to a more concise description in terms of graphs colored by spins j, these are spin networks.

3.4 $SU(2)$ spin networks

This section introduces the main objects of study in this thesis, $SU(2)$ spin networks. The focus will be on introducing the graphical tools specific to this context; in the next two chapters we will take up the study of the semiclassical geometries associated to spin networks. The most significant difference between the networks discussed in this section and those of the previous section on Euclidean invariants is that the ambient space will be taken to be $\mathbb{C}^2$, this means that the ϵ tensor will only carry two indices ϵ_{AB} or $\epsilon^{AB}, (A, B = 1, 2)$. The choice of this space is due to the fact that we will be studying $SU(2)$ and interested in its action on two component spinors $\psi^A = (z_1, z_2)$.

3.4.1 String diagrams

Spin networks evolved out of the string diagrams considered in 3.2. They can be seen as a special case of string diagrams that has been adapted to the structure of $SU(2)$. The fundamental string diagrams for the ambient space $\mathbb{C}^2$ are,

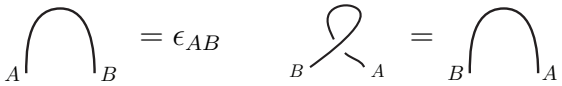


(3.69)

Note that we have chosen not to represent the ϵ tensor with any sort of node but instead by a turn in our string — this has several impacts on the diagram rules. As discussed in section 3.2, for a graphical calculus planar isotopy is king. To retain planar isotopy with the strings of (3.69) it is necessary to introduce two auxiliary rules. For example, we once more want to freely add loops to our string,

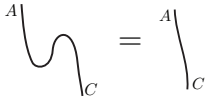
(3.70)

This is only consistent with the antisymmetry of the ϵ tensor if we associate a minus sign to all crossings,[5]



(3.71)

Similarly we want,

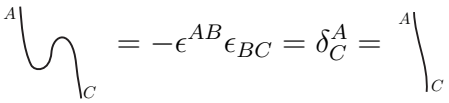


(3.72)

and algebraically we find,

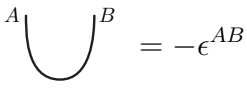
$$\epsilon^{AB}\epsilon_{BC} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\delta^A_C. \tag{3.73}$$

To bring the two equations into a nice correspondence we associate a minus sign to all local minima in a diagram, with the convention that extrema are measured with respect to vertical on the page. Then we have,



(3.74)

This last rule is succinctly summarized by adding the diagram



(3.75)

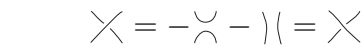
[5]A point of clarification: notice that, if read from left to right, the order of the indices of the twisted graph and those of the ϵ tensor are no longer the same. As we will see momentarily crossings can be removed by reduction but if there is any confusion about ordering of indices due to crossing just view the crossing as two Kronecker deltas and the correspondence is clear:

to our list of fundamental diagrams (3.69).

These conventions take a moment to internalize but allow the retention of the full machinery of the Reidemeister moves and are invaluable for this reason. As an example of the sort of thing one becomes accustomed to, the dimensionality graph ○ yields a negative result:

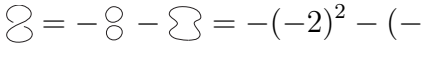
$$= -\epsilon^{AB}\epsilon_{AB} = -\delta^A_A = -2. \tag{3.76}$$

Just as in the earlier $SO(3)$ case, a basic reduction rule is essential. In two dimensions the ϵ tensor is quadratically dependent on the Kronecker delta and it is simply this dependency that gives the reduction rule:



$$= -\epsilon^{AB}\epsilon_{CD} - \delta^A_C\delta^B_D = -\delta^A_D\delta^B_C. \tag{3.77}$$

A nice example is ⊗, which we now have two ways to calculate, by isotopy this is ⊗ = ○ = −2, and by reduction it is,



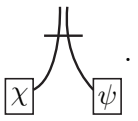
$$= -(-2)^2 - (-2) = -2. \tag{3.78}$$

This exhibits how simple the graphical calculus is in this context. A closed network corresponds to a number, and that number is computed by repeated application of the reduction rule and the dimensionality graph, ○ = −2.

Perhaps it is unsurprising that this framework becomes much richer when you introduce composite spin systems. The richness of the full machinery is already indicated in the addition of two spin-1/2 systems. A general state of the tensor product space $\mathcal{H}_{j_1} \otimes \mathcal{H}_{j_2} = \mathcal{H}_{1/2} \otimes \mathcal{H}_{1/2}$ can be decomposed into its antisymmetric and symmetric parts,

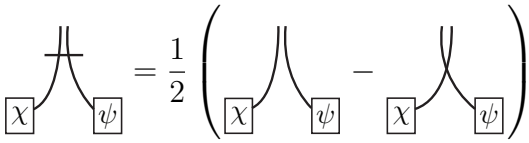
$$\phi^{AB} = \phi\epsilon^{AB} + \phi^{(AB)}, \tag{3.79}$$

with $\phi = \frac{1}{2}\epsilon_{CD}\phi^{CD}$. This decomposition corresponds to the break up of the total space into its rotationally invariant, $j = 0$, singlet component, which explicitly appears as the invariant ϵ tensor in (3.79), and the triplet, $j = 1$, state $\phi^{(AB)}$. An essential property of a graphical calculus for $SU(2)$ networks is a good representation of this decomposition into angular momentum subspaces. We now begin to specialize our conventions to this context. Consider the following graph connecting two spinors $\chi^A, \psi^A \in \mathcal{H}_{1/2}$:



. (3.80)

The bar indicates an antisymmetric sum of *graphs*,



$$= \frac{1}{2}\left(\quad - \quad \right) = \frac{1}{2}\left(\chi^A\psi^B + \chi^B\psi^A\right), \tag{3.81}$$

notice that this is a symmetrization of *indices* due to the rule about crossings in graphs! In fact, as we will see momentarily we can always decompose $\phi^{(AB)}$ into such a symmetrized product of spinors $\phi^{(AB)} = \chi^{(A}\psi^{B)}$. The antisymmetrization of two strings picks out precisely the $j = 1$ component of the tensor product. This works not just for $j = 1$ but for arbitrary j.

The idea is to extend the decomposition of (3.77) to spinorial tensors of general type (r, s), henceforth we also refer to these tensors as spinors.[6] These tensors are elements of $\mathcal{S} \otimes \cdots \otimes \mathcal{S} \otimes \mathcal{S}^* \otimes \cdots \otimes \mathcal{S}^*$, introducing the shorthand $\mathcal{S} \equiv \mathbb{C}^2 = \mathcal{H}_{1/2}$ for spinor space and $\mathcal{S}^*$ for its dual. Instead of assuming a metrical structure we can use the ϵ-tensor to raise and lower indices, e.g. for $\chi^A \in \mathcal{S}$ we have,

$$\chi_B = \chi^A \epsilon_{AB}, \tag{3.82}$$

and this map provides an identification of $\mathcal{S}$ with $\mathcal{S}^*$.[7] Now, assume that $\Phi_{\ldots AB\ldots} = \Phi_{\ldots[AB]\ldots}$, where the ellipsis denote further indices that are not essential to what follows. The pair of indices AB can be replaced by ϵ_{AB} using the reduction rule (3.77),

$$\epsilon_{AB}\epsilon^{CD}\Phi_{\ldots CD\ldots} = (\delta^C_A\delta^D_B - \delta^D_A\delta^C_B)\Phi_{\ldots CD\ldots} = 2\Phi_{\ldots AB\ldots}, \tag{3.84}$$

and so,

$$\Phi_{\ldots[AB]\ldots} = \frac{1}{2}\epsilon_{AB}\Phi_{\ldots C}{}^{C}{}_{\ldots}. \tag{3.85}$$

Thus a general spinor can be decomposed into a sum of pieces in which all antisymmetry is expressed explicitly in terms of ϵ-tensors. This classification can be extended further by considering the completely symmetric groups of indices. Assuming $\Theta_{A\ldots B} = \Theta_{(A\ldots B)}$, introduce a basis into $\mathcal{S}$ (with the dual basis of $\mathcal{S}^*$) and form the complete contraction with $\xi^A = (z, 1)$: $\Theta_{A\ldots B}\xi^A \cdots \xi^B$. This is a polynomial in z and factorizes into linear factors $\chi_A\xi^A \cdots \psi_B\xi^B$ and so,

$$\Theta_{(A\ldots B)} = \chi_{(A} \cdots \psi_{B)}. \tag{3.86}$$

The spinors $\chi^A, \ldots, \psi^B$ are only defined up to an overall scale because of the same ambiguity in the factorization of the polynomial. The results (3.85) and (3.86) provide a powerful classification of spinors.

Returning to $SU(2)$ we see that the space of completely symmetric spinors transforms into itself under an $SU(2)$ transformation (due to (3.86)). Thus it forms a

[6] This presentation follows [78].

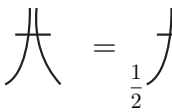
[7] There is an important convention built into the ordering of indices in this equation, which as usual reflects the need to be vigilant about the antisymmetry of the ϵ-tensor, so that,

$$\epsilon_{AB}\chi^B = -\chi^B\epsilon_{BA} = -\chi_A. \tag{3.83}$$

A mnemonic that works for both this and the dual identification, $\alpha^A = \epsilon^{AB}\alpha_B$, is "down and to the right."

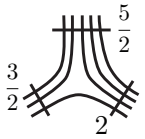
representation, one that is in fact irreducible. If the spinor carries n indices this is the spin-j representation $\mathcal{H}_j$ with $j = n/2$, exactly as would be expected for the constructive addition of n spin-1/2 particles, and has dimension $(2j + 1)$. We'd like to capture these observations graphically.

The core of the graphical example illustrating the addition of two spin-1/2 systems above, (3.80), was the invariant tensor,



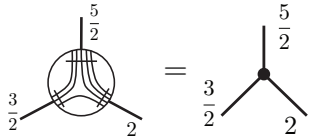
(3.87)

where on the right the spin representation is indicated explicitly. This diagram represents a tensor that can be used to couple two spin-1/2 systems to yield the triplet, spin-1 state (3.81). As we have seen antisymmetrization of strings corresponds to symmetrization of indices and the index symmetric subspaces carry the irreps of $SU(2)$, so this basic strategy can be extended to any number of strings,[8]



(3.88)

This gives us a way to build up invariant tensors that couple two given $SU(2)$ irreps into a third; these diagrams are like Clebsch-Gordan coefficients. More precisely they are one way of writing Wigner's $3j$-symbol, a symmetrical version of the Clebsch-Gordan coefficients. The string diagram notation is useful because of its explicitness and because it represents a direct computational tool, called chromatic evaluation, in which closed diagrams can be expanded into sums of loops and evaluated to numbers as at (3.78). However, it is clear that for even moderately large angular momentum j, it is no longer practical to write out the strings. Instead treat the bars as ties which bind the strings into a rope and label the rope by the appropriate j,



(3.89)

This succinct diagram is our first example of a graphical spin network. Before leaving behind the string diagrams to discuss spin networks more axiomatically we present a simple derivation of the Clebsch-Gordan conditions facilitated by the string language.

Consider the coupling of two angular momenta j_1 and j_2. In the string diagram language this is a routing of $a = 2j_1$ and $b = 2j_2$ strands through a node resulting in

[8] As with the symmetrization and antisymmetrization of indices the bar notation carries a normalization factor $1/n! = 1/(2j)!$.

a rope with $c = 2j_3$ strands, a, b and c all integers. The result of the coupling depends on how many of the a-strands are connected to b-strands. If no a- and b-strands are connected then $c = a + b$ and $j_3 = j_1 + j_2$. If 1 a-strand is connected to 1 b-strand then $c = a + b - 2$ and $j_3 = j_1 + j_2 - 1$. Continuing to connect strands one at a time j_3 decreases in integer steps until all possible ab-strand connections are made, the remaining strands yield the minimum values $c = |a - b|$ and $j_3 = |j_1 - j_2|$. Because each string has two ends $a + b + c$ is even and it follows that the sum $j_1 + j_2 + j_3$ is integer! So, it is possible to obtain j_3 in the coupling of j_1 and j_2 if

$$|j_1 - j_2| \leq j_3 \leq j_1 + j_2 \qquad \text{and} \qquad j_1 + j_2 + j_3 = n, \tag{3.90}$$

with n an integer. These are the Clebsch-Gordan conditions and this method of obtaining them is refreshing simple and intuitive. These conditions also arise in the semiclassics of the $3j$-symbol in a beautiful manner: the Lagrangian manifold of the $3j$-symbol is topologically non-trivial (but not a torus!) and the Bohr-Sommerfeld quantization of one of its non-trivial basis contours gives the integer perimeter condition just described [37].

3.4.2 Spin networks

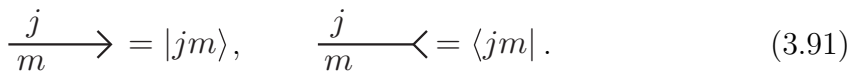
In this section we provide an overview of the correspondence between the quantum theory of angular momentum and the graphical spin network representation.[9] There is unfortunately no universal notation for spin networks, however the variations are small and correspondences can usually be made quickly. Below we set our conventions and explain some subtleties that can be confusing. There are a number of references [99, 100, 76, 101, 102, 103, 104] and we particularly recommend Moussouris' thesis as an introduction [105], Yutsis et al for the recoupling perspective [106] and Stedman for an overview [91].

There are both similarities and differences between the spin network notation and that of the string diagrams from the last section. Like string diagrams the links of a spin network can be arbitrarily deformed but unlike string diagrams there is no reduction rule applying at the overlap of two links. Nodes of a spin network will be marked by a small dot and all other crossings have no significance. With spin networks we drop any attempt to use orientation on the page to track index position, this will be replaced by a decoration on the graph. One complication in the use of spin networks is tracking phases, but the graphical notation is no worse than algebraic manipulations; phases can be a pain. We adopt the standard physics conventions for the phases of the $|jm\rangle$ states, [107].

Appropriating Stedman's convention we will represent bra and ket vectors by wide

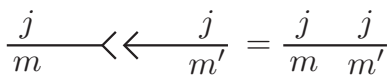
[9] Many of the figures of this section are courtesy of Robert Littlejohn and parts of the exposition follow [98].

angle arrows or chevrons,



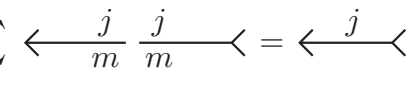
(3.91)

The arrow head of the ket points away from the line and that of the bra points toward the line. Orthonormality of these basis vectors, $\langle jm'|jm\rangle = \delta_m^{m'}$, looks like



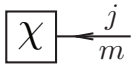
(3.92)

and completeness, $\sum_m |jm\rangle\langle jm| = \mathbb{1}_j$, is a flighted arrow,



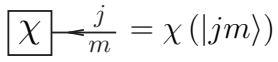
(3.93)

As before more general tensor are represented with labelled nodes,



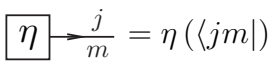
(3.94)

and the small arrow indicates that χ transforms covariantly with respect to ($SU(2)$) rotations,



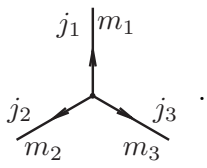
(3.95)

The intermediate equality displays χ acting on the $|jm\rangle$ basis in the tensor notation of the invariant theory section. Similarly a contravariant vector η is represented by



(3.96)

The trivalent tensor introduced at the end of the last section, which is invariant, is called a $3j$-symbol and denoted,

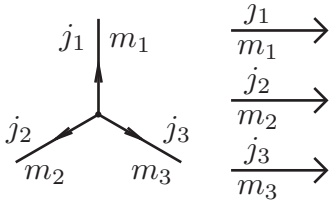
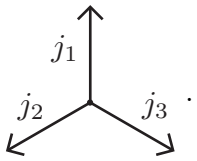


(3.97)

This tensor couples the states of the three representations $\mathcal{H}_{j_1}$, $\mathcal{H}_{j_2}$ and $\mathcal{H}_{j_3}$ to a state contained in the rotationally invariant subspace $\mathcal{K} \subset \mathcal{H}_{j_1} \otimes \mathcal{H}_{j_2} \otimes \mathcal{H}_{j_3}$. This is the heart of the intertwiner definition briefly mentioned in the introduction to

this chapter. To satisfy the Clebsch-Gordan conditions we are assuming that the triple (j_1, j_2, j_3) is triangular and has integer sum. Each irrep in the Clebsch-Gordan series occurs only once and so the rotationally invariant subspace, with $j = 0$, is one dimensional $\dim \mathcal{Z} = 1$. This means that the rotationally invariant subspace is spanned by a single vector and in honor of Wigner we call this the Wigner state $|W\rangle$:

$$|W\rangle = \sum_{m\text{'s}} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} |j_1 m_1\rangle |j_2 m_2\rangle |j_3 m_3\rangle$$

$$= \sum_{m\text{'s}}$$  $$=$$  . (3.98)

The first line of this equation introduces the standard notation for the $3j$-symbol,

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}. \tag{3.99}$$

Henceforth we identify the symbol and its network and refer to them interchangeably.

The semiclassics of the $3j$-symbol and the Wigner state were studied at length in our work [37]. Remarkably the Lagrangian manifold associated to this state is not a torus, this is in spite of the fact that the Schwinger model can be used to associate a classically integrable system to the Wigner state. This is possible because the components of total angular momentum commute on the level set where the total angular momentum vanishes despite their not commuting in general. This is a general feature of the semiclassics of closed spin networks built out of the $3j$-symbols.

The $3j$-symbol is the fundamental unit of the spin network diagrams in the same sense that the ϵs and δs were for the general linear group. A closed spin network is one with no dangling links, i.e. a complete contraction of $3j$-symbols. This means that all $SU(2)$ invariants (without consideration of additional tensors of physical interest) are represented by closed trivalent graphs colored by spins j, these are the $3nj$-symbols. Generally $3n$ gives the number of links of the network but the case $n = 1$ is exceptional; the theta graph $\ominus$ is the normalization of the Wigner state and in Wigner's conventions is taken to be one, $\langle W|W\rangle = 1$. For $n > 2$ there are multiple distinct networks with $3n$ links and these are called the types of the $3nj$-symbol in the conventions of Yutsis et al [106]. In this dissertation we only treat symbols of type I although other types will almost certainly end up being of interest.

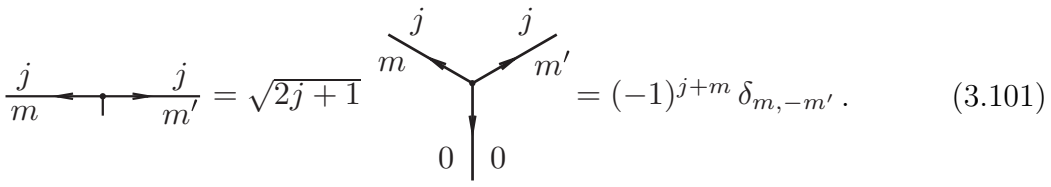
While the trivalent graphs are of primary interest in relation to the $3nj$-symbols, in quantum gravity we will generally be interested in graphs containing higher valent nodes. The reason for this, as will be discussed in the next chapter, is that the volume operator vanishes for trivalent nodes and so a graph is only associated to a

space with extension if it has higher valency. When more than three irreps of $SU(2)$ are combined, or in other words, when we couple more than three angular momenta, the rotationally invariant subspace $\mathcal{K}$ is no longer one dimensional. We call this rotationally invariant subspace the intertwiner space and formally it is,

$$\mathcal{K}_n = \mathrm{Inv}\left(\mathcal{H}_{j_1} \otimes \cdots \otimes \mathcal{H}_{j_r}\right), \tag{3.100}$$

where we are assuming that r links meet at the node n. The volume operator provides one way to label its basis states, more on this point in the next section.

Only a few more clarifications and definitions remain. It turns out to be useful to introduce a special case of the $3j$-symbol where one of the js is taken to be zero. Along with Stedman we call this a $2j$-symbol and denote it with a stub where the third link once was,



(3.101)

Because the total angular momentum associated to a $3j$-symbol is zero the two js of a $2j$-symbol must be equal and have opposite m quantum numbers. Additionally the phase $(-1)^{j+m}$ and normalization come from evaluating,

$$\begin{pmatrix} j & j \\ m & m' \end{pmatrix} = \begin{pmatrix} j & 0 & j \\ m & 0 & m' \end{pmatrix} = \frac{(-1)^{j+m}\delta^{m,-m'}}{\sqrt{2j+1}}. \tag{3.102}$$

The importance of the $2j$-symbol is that it gives a second way to raise and lower indices that is different from the one associated with the metric $\hat{G}^j$ of the Hilbert space $\mathcal{H}_j$. We have already seen this in the context of the string diagram, the ϵ_{AB} tensor was used to lower indices instead of a metric on $\mathcal{S}$.

The metric $\hat{G}^j$ is the usual map of quantum mechanics $\hat{G}^j : \mathcal{H}_j \to \mathcal{H}_j^*$, given by Hermitian conjugation,

$$\hat{G}^j\left(|\psi\rangle\right) = \langle\psi| \tag{3.103}$$

and is antilinear. However we can also consider a map $\hat{K} = \hat{G}^j\hat{\Theta} : \mathcal{H}_j \to \mathcal{H}_j^*$, here $\hat{\Theta}$ is the time reversal operator on angular momenta, it acts on basis vectors by,

$$\hat{\Theta}|jm\rangle = (-1)^{j-m}|j,-m\rangle, \tag{3.104}$$

and is also antilinear (see e.g. Sakurai [108], which has the same phase conventions). The $\hat{K}$ map action on $|\psi\rangle$ is,

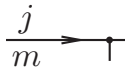
$$\hat{K}|\psi\rangle = \langle\psi|\hat{\Theta}^\dagger, \tag{3.105}$$

the time reversal of the bra $\langle\psi|$ ($\Theta^\dagger$ acts from the right). On basis vectors this is

$$\hat{K}|jm\rangle = (-1)^{j-m}\langle j,-m| \tag{3.106}$$

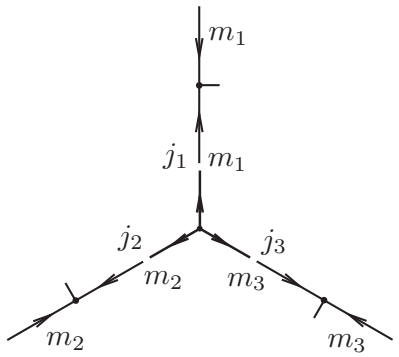
and we see that the $2j$-symbol is the component form of this map. Note that $\hat{K}$ is linear, not anitlinear, and the component matrix is unitary not symmetric or Hermitian.

Indices of spin networks are generally raised and lowered using the $2j$-symbols (3.101) and

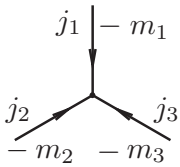


$$= (-1)^{j+m}\,\delta_{m,-m'}\,. \tag{3.107}$$

For example reversing the arrows of the $3j$-symbol is achieved by contracting with three $2j$-symbols

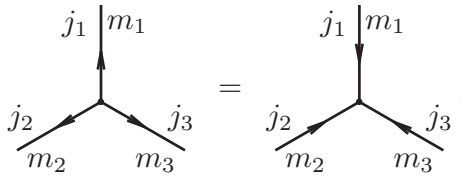


$$= (-1)^{j_1+j_2+j_3+m_1+m_2+m_3}$$



. (3.108)

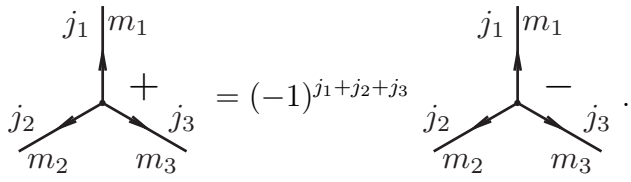
and this can be related to the original contravariant $3j$-symbol as follows: Because the total angular momentum is zero the sum of the ms is zero. Reversing all of the signs of the ms is a symmetry of the $3j$-symbol which incurs the phase $(-1)^{j_1+j_2+j_3}$. Finally, the sum of the js is an integer and so the total phase shift is $(-1)^{2(j_1+j_2+j_3)} = 1$ and we find,



. (3.109)

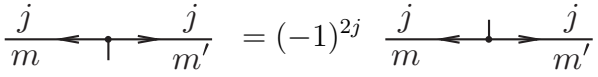
The $3j$-symbol node is not symmetric with respect to interchange of the legs of the network. For this reason it is necessary to indicate the order in which the labels are to be read. A + sign at the node indicates that the indices should be read in the counterclockwise direction around the node, while as minus sign indicates clockwise. This is analogous to the positive angle convention in the polar plane. If a node bears

no decoration we conventionally assume that it is positively oriented. Interchange of the m_1 and m_2 legs of the $3j$-symbol results in the phase $(-1)^{j_1+j_2+j_3}$,



(3.110)

It is because of these phases that the stub on the $2j$-symbol is necessary



(3.111)

(again there's an implicit + at the nodes).

That is it. Any spin network can be built out of these basic ingredients and the symbolic representation of a network can be recovered, including all of the phases.

3.5 Spin networks and quantum gravity

In this short section we motivate why spin networks are an interesting basis for the Hilbert space $\mathcal{H}$ of loop quantum gravity and briefly indicate how this result is obtained.[10] Recall from the introduction that $\mathcal{H}$ is built out of the graph Hilbert spaces $\mathcal{H}_\Gamma = L^2\left(SU(2)^L/SU(2)^N\right)/\sim$. We begin by considering the non-gauge invariant link Hilbert space $L^2(SU(2)^L)$. There are natural derivative operators on $L^2(SU(2))$ that act as the "momenta" of loop gravity. Let $\tau_i = -i/2\sigma_i$, $(i = 1, 2, 3)$ be a basis of the Lie algebra of $SU(2)$ where the σ_i are the Pauli matrices. Then these derivative operators are the left invariant vector fields,

$$J_i\psi(h) \equiv i\frac{d}{dt}\psi\left(he^{i\tau_i}\right)\Big|_{t=0} \qquad (i = 1, 2, 3) \tag{3.112}$$

(in this section we drop the hats on operators). More generally such an operator can be associated to each link ℓ of Γ and we denote these operators $\boldsymbol{J}_\ell$. The gauge invariance (1.1) implies that

$$\sum_{\ell\in n} \boldsymbol{J}_\ell = 0 \tag{3.113}$$

at each node n of Γ. We will have more to say about this condition in the next section. In his seminal work on spin networks Penrose introduced the gauge invariant operator

$$G_{\ell\ell'} = \boldsymbol{J}_\ell \cdot \boldsymbol{J}_{\ell'} \tag{3.114}$$

[10] Again this section closely follows [7]

with ℓ and ℓ' both initiating at the same node $s(\ell) = s(\ell')$. Penrose showed that in the semiclassical limit this operator can be interpreted as defining angles in three dimensional space and he realized it was effectively a quantum metric. In the next chapter we show that in this limit the nodes of the graph can be interpreted as polyhedra and that the operators $\boldsymbol{J}_\ell$ can be associated to classical area vectors $\boldsymbol{A}_\ell$. These area vectors are normal to the faces of the polyhedron and have magnitudes $A_\ell = |\boldsymbol{A}_\ell|$ equal to the polyhedron's face area. Thus $G_{\ell\ell'}$ acts like a quantum metric with diagonal components the face areas squared and off-diagonal components

$$G_{\ell\ell'} \sim A_\ell A_{\ell'} \cos\left(\theta_{\ell\ell'}\right) \tag{3.115}$$

here $\theta_{\ell\ell'}$ is the dihedral angle between the faces. We use the $\sim$ here to indicate that this only makes sense semiclassically.

This picture also gives rise to a natural geometrical proposal for the squared volume of a quantum tetrahedron

$$V = \frac{\sqrt{2}}{3}\sqrt{|\boldsymbol{J}_{\ell_1} \cdot (\boldsymbol{J}_{\ell_2} \times \boldsymbol{J}_{\ell_3})|} \tag{3.116}$$

which is modeled on the classical relation

$$V^2 = \frac{2}{9}\left|\boldsymbol{A}_{\ell_1} \cdot (\boldsymbol{A}_{\ell_2} \times \boldsymbol{A}_{\ell_3})\right|. \tag{3.117}$$

This volume operator is the main object of study in the next chapter. The great advantage of the spin network basis is that the geometrical operators $G_{\ell\ell} = \boldsymbol{J}_\ell \boldsymbol{J}_\ell$ and Q are diagonal in this basis. A note of caution, tetrahedra only being associated to 4-valent nodes, we cannot say that the volume operator is diagonalized in general. In fact, the study of the volume for higher valence is an interesting current area of research, see the end of Chapter 4 for some partial results.

Now, for the more technical part. According to the Peter-Weyl theorem $L^2\left(SU(2)^L\right)$ can be decomposed into irreducible representations

$$L^2\left(SU(2)^L\right) = \bigoplus_{j_\ell} \bigotimes_{\ell} \left(\mathcal{H}^*_{j_\ell} \otimes \mathcal{H}_{j_\ell}\right). \tag{3.118}$$

Once again recalling the gauge transformation,

$$\psi(h_\ell) \mapsto \psi(g_{s(\ell)} h_\ell g^{-1}_{t(\ell)}), \tag{3.119}$$

note that the Hilbert space factors above are naturally associated with the source and target nodes. So instead of collecting terms over links we do it over the nodes

$$L^2\left(SU(2)^L\right) = \bigoplus_{j_\ell} \bigotimes_{n} \mathcal{H}_n, \tag{3.120}$$

with

$$\mathcal{H}_n = \left(\bigotimes_{\ell \in s(n)} \mathcal{H}_\ell^* \right) \otimes \left(\bigotimes_{\ell \in t(n)} \mathcal{H}_\ell \right). \tag{3.121}$$

So far we have been considering the space $L^2\left(SU(2)^L\right)$ but we are really interested in gauge invariant states and we can also build these up in a piecewise manner. Restrict attention to the gauge invariant part of the node Hilbert space

$$\mathcal{K}_n = \mathrm{Inv}\left(\mathcal{H}_n\right). \tag{3.122}$$

The nodes are gauge invariant under the diagonal action of $SU(2)$, that is to say, they are subjected to the condition (3.113). But this is precisely what we require at the nodes of a spin network. The volume operator acts on these intertwiner spaces $\mathcal{K}_n$ and we can use its eigenvalues v_n to label a choice of basis. Finally, we get the graph space by putting together all the node pieces

$$\mathcal{H}_\Gamma = L^2\left(SU(2)^L/SU(2)^N\right) = \bigoplus_{j_\ell} \bigotimes_n \mathcal{K}_n. \tag{3.123}$$

In summary, the geometrical spin network basis states $|\Gamma\ j_\ell\ v_n\rangle$ are determined by an abstract graph Γ, colored by spins j_ℓ, which determine the areas between neighboring regions of space, and a specification of basis in the intertwiner spaces, which determines the volume of the regions represented by the nodes of the graph Γ.

In the next chapter we develop the polyhedral picture for the nodes and compile evidence supporting the definition (3.116) for the quantum volume operator of a tetrahedral piece of space.

Part II

Applications

Chapter 4

The Volume of Space

4.1 Introduction

In this chapter we take up the quantization of space.[1] Section 3.5 and the introduction gave the qualitative picture: in loop gravity quantum states are built upon a spin colored graph. The nodes of this graph represent three dimensional grains of space and the links of the graph encode the adjacency of these regions. We will be interested in the semiclassical description of the grains of space. In particular, following Bianchi, Doná and Speziale [109] we will interpret these grains as giving rise to convex polyhedra in this limit. This geometrical picture suggests a natural proposal for the quantum volume operator of a grain of space: it should be an operator that corresponds to the volume of the associated classical polyhedron. The main body of the chapter provides a detailed analysis of such an operator, for a single 4-valent node of the graph. The valency of the node determines the number of faces of the polyhedron and so we will be focusing on classical and quantum tetrahedra. Remarkably, the space of convex polyhedra with fixed face areas has a natural phase space and symplectic structure. This was first shown by Kapovich and Millson [110] in a different context. Using this kinematics we study the classical volume operator and perform a Bohr-Sommerfeld quantization of its spectrum. This quantization is in good agreement with previous studies of the volume operator spectrum [111, 112, 113, 114, 115, 116, 117, 118, 119] and provides a simplified derivation.

The plan of the chapter is as follows: We begin by recalling and summarizing the setup for the node Hilbert Spaces $\mathcal{K}_n$ for general valency. Next we describe how this space limits to a classical phase space and describe its Poisson structure. The volume operator has been studied extensively in loop gravity. In section 4.3 we connect with this literature and briefly describe some difficulties arising in these approaches. The generalities give way to a discussion of what is known for the quantum tetrahedron

[1] This chapter is based on joint work with Eugenio Bianchi to appear in Physical Review Letters [71] and a longer paper in preparation. A few passages are adapted from these works.

in section 4.3.1. In sections 4.4 and 4.5 we describe a phase space for the classical tetrahedron and take up its Bohr-Sommerfeld quantization. Elliptic functions play a prominent role in this quantization and section 4.6 develops tools for analyzing the largest and smallest volume eigenvalues using the properties of elliptic functions. Wavefunctions for the volume operator are constructed in section 4.7. To the best of our knowledge this is the first time that analytic formulas for these wavefunctions have appeared. In section 4.8 we discuss partially completed research into the extension of this work to triangular prisms. Section 4.9 summarizes our progress and provides outlook for future studies of the volume operator.

4.2 Setup

To begin we collect the definitions used in setting up the loop gravity Hilbert space and connect them to a classical theory. As discussed at the end of the last chapter the graph Hilbert space $\mathcal{H}_\Gamma$ can be built out of gauge-invariant Hilbert spaces at the nodes. For this entire chapter we will focus on a single node n and its Hilbert space $\mathcal{K}_n$. Let us assume that the node is F-valent and call the spins labeling its links j_r $(r = 1, \ldots, F)$. To each representation j_r we associate the vector space $\mathcal{H}_{j_r}$. This is the vector space on which the $SU(2)$ generators $\boldsymbol{J}_r$ act. The standard basis $|j_r m_r\rangle$ is labelled by eigenvalues of the Casimir $J_r^2 = \boldsymbol{J}_r \cdot \boldsymbol{J}_r$ and the component of $\boldsymbol{J}_r$ in a given direction $\boldsymbol{n}$, i.e. $\boldsymbol{n} \cdot \boldsymbol{J}_r$. The space $\mathcal{K}_n$ is defined as the subspace of the tensor product $\mathcal{H}_{j_1} \otimes \cdots \otimes \mathcal{H}_{j_F}$ that is invariant under global $SU(2)$ transformations (the diagonal action)

$$\mathcal{K}_n = \mathrm{Inv}\left(\mathcal{H}_{j_1} \otimes \cdots \otimes \mathcal{H}_{j_F}\right). \tag{4.1}$$

We call $\mathcal{K}_n$ the space of intertwiners. The diagonal action is generated by the operator $\boldsymbol{J}$,

$$\boldsymbol{J} = \sum_{r=1}^{F} \boldsymbol{J}_r. \tag{4.2}$$

States of $\mathcal{K}_n$ are called intertwiners and can be expanded in the $|j_r m_r\rangle$ basis described above, for $|i\rangle \in \mathcal{K}_n$,

$$|i\rangle = \sum_{m\text{'s}} i^{m_1 \cdots m_F} |j_1 m_1\rangle \cdots |j_F m_F\rangle. \tag{4.3}$$

The components $i^{m_1 \cdots m_F}$ transform as a tensor under $SU(2)$ transformations in such a way that the condition

$$\boldsymbol{J}|i\rangle = 0 \tag{4.4}$$

is satisfied. These are precisely the invariant tensors captured graphically by spin networks.

The finite dimensional intertwiner space $\mathcal{K}_n$ can be understood as the quantization of a classical phase space. We setup this correspondence in three steps. We begin on

the classical side with a classic result due to Minkowski [120]. The importance of this result for loop gravity was first appreciated by Eugenio Bianchi.

Minkowski's theorem state the following: given F vectors $\boldsymbol{A}_r \in \mathbb{R}^3$ $(r = 1, \ldots, F)$ whose sum is zero

$$\boldsymbol{A}_1 + \cdots \boldsymbol{A}_F = 0 \tag{4.5}$$

then up to rotations and translations there exists a unique convex polyhedron with F faces associated to these vectors. Furthermore, the vectors $\boldsymbol{A}_r$ can be interpreted as the outward pointing normals to the polyhedron and their magnitudes $A_r = |\boldsymbol{A}_r|$ as the face areas. Minkowski's proof is not constructive and so this is strictly an existence and uniqueness theorem. We call the process of building one of these polyhedra (given the area vectors) the Minkowski reconstruction problem. In the case of the tetrahedron the reconstruction is trivial but for larger F this is a difficult problem [109]. One of our results in section 4.8 is an analytical solution of the reconstruction in the case of 5 faces. These convex polyhedra are our semiclassical interpretation of the loop gravity grains of space.

The second step of the construction is to associate a classical phase space to these polyhedra. Following Kapovitch and Millson, we interpret the partial sums

$$\mu_k \equiv \left| \sum_{r=1}^{k+1} \boldsymbol{A}_r \right| \qquad\qquad (k = 1, \ldots, F-3) \tag{4.6}$$

as generators of rotations about the $\boldsymbol{\mu}_k = \boldsymbol{A}_1 + \cdots + \boldsymbol{A}_{k+1}$ axis. This interpretation follows naturally from considering each of the $\boldsymbol{A}_r$ vectors to be a classical angular momentum, so that really $\boldsymbol{A}_r \in \Lambda_r \cong \mathbb{R}^3$. Recall that Λ is the name that we gave angular momentum space in Chapter 2. As discussed in Chapter 2 there is a natural Poisson structure on each Λ_r and this extends to the bracket

$$\{f, g\} = \sum_{r=1}^{F} \boldsymbol{A}_r \cdot \left(\frac{\partial f}{\partial \boldsymbol{A}_r} \times \frac{\partial g}{\partial \boldsymbol{A}_r} \right) \tag{4.7}$$

on F copies of Λ, here f and g are arbitrary functions of the $\boldsymbol{A}_r$. With this Poisson bracket the μ_k do in fact generate rotations about the axis $\boldsymbol{A}_1 + \cdots \boldsymbol{A}_{k+1}$. This geometrical interpretation suggests a natural conjugate coordinate, namely the angle of the rotation. Let ϕ_k be the angle between the vectors

$$\boldsymbol{v}_k = \left(\sum_{r=1}^{k} \boldsymbol{A}_r \right) \times \boldsymbol{A}_{k+1} \qquad \text{and} \qquad \boldsymbol{w}_k = \left(\sum_{r=1}^{k+1} \boldsymbol{A}_r \right) \times \boldsymbol{A}_{k+2}. \tag{4.8}$$

It is an easy check to show that

$$\{\mu_k, \phi_l\} = \delta_{kl}. \tag{4.9}$$

The pairs (μ_k, ϕ_k) are canonical coordinates for a classical phase space of dimension $2(F-3)$. We call this the space of shapes and denote it $\mathcal{P}(A_1, \ldots, A_F)$ or more briefly $\mathcal{P}_F$.

The explicit inclusion of the A_r in $\mathcal{P}(A_1, \ldots, A_F)$ highlights a parametric dependence of the space of shapes on the magnitudes A_r. We explain this by adopting another perspective on $\mathcal{P}_F$, viewing it as a symplectic reduction of the product of F two spheres, $(S^2)^F$. Upon fixing the magnitude A_r the angular momentum vector $\boldsymbol{A}_r$ is restricted to a sphere of radius A_r in Λ_r. This sphere is a symplectic leaf of the Poisson manifold Λ_r and the collection of all the spheres $(S^2)^F$ can be endowed with the product symplectic structure. If we now symplectically reduce $(S^2)^F$ by the zero level set of the momentum map

$$\boldsymbol{A} = \sum_{r=1}^{F} \boldsymbol{A}_r \tag{4.10}$$

we once again obtain $\mathcal{P}_F$. This perspective also explains the dimension of the reduced space; because zero is a fixed point of the group action we lose twice as many dimensions as there are components of the momentum map, $\dim \mathcal{P}_F = 2F - 6$.

In our third and final step we indicate a route to showing that that quantization of $\mathcal{P}_F$ is the Hilbert space $\mathcal{K}_n$ of an F-valent node n. The F angular momentum vectors $\boldsymbol{A}_r$ can be lifted via the Schwinger map to the phase space of $2F$ harmonic oscillators, $\mathbb{C}^{2F}$. These oscillators can be quantized in the standard fashion and once again the Schwinger map can be used to reduce to F quantum angular momenta, on the quantum side. The constraint

$$\boldsymbol{A}_1 + \cdots + \boldsymbol{A}_F = 0 \tag{4.11}$$

becomes

$$\boldsymbol{J}_1 + \cdots \boldsymbol{J}_F = 0 \tag{4.12}$$

which is precisely the gauge invariance condition 3.113. Perhaps the content of this paragraph is obvious, but no one ever explains where the classical vector model of angular momentum originates. Before specializing to the case $F = 4$ we review some of the loop gravity literature on the volume operator.

4.3 Volume operators in loop gravity

Most of the loop gravity research on the volume operator has been done in the context of the original canonical quantization of general relativity. To make contact with this work we briefly review the variables of the canonical framework. In Ashtekar's formulation of classical general relativity the gravitational field is described in terms of the triad variables $\boldsymbol{E}$.[2] This triad corresponds to a 3-metric h and is called the

[2]Technically $\boldsymbol{E}$ is an inverse densitized triad.

electric field. In distinction to our discussion up to now, the original formulation considered spin network graphs Γ to be embedded in a three manifold Σ, interpreted as space. The elementary quantum operator that measures the geometry of space corresponds to the flux of the electric field through a surface S. When such a surface is punctured by a link of the spin network graph Γ the flux can be parallel transported, back along the link, to the node using the second of Ashtekar's variables, the Ashtekar-Barbero connection. This results in an $SU(2)$ operator that acts on the intertwiner space $\mathcal{K}_n$ at the node n. We label the F links at the node with the index $r = 1, \ldots, F$ and call this operator $\boldsymbol{E}_r$. The parallel-transported flux operator at the node is proportional to the generator of $SU(2)$ transformations introduced above

$$\boldsymbol{E}_r = 8\pi\gamma\ell_{\mathrm{Pl}}^2 \boldsymbol{J}_r, \tag{4.13}$$

where γ is a free parameter of the theory called the Barbero-Immirzi parameter and ℓ_{Pl} is the Planck length. Here we have used the embedded picture to introduce these operators but they can just as well be defined (starting with the $\boldsymbol{J}_r$) this way in the combinatorial picture described in previous sections.

The volume of a region of space R is obtained by regularizing and quantizing the classical expression

$$V = \int_R d^3x\sqrt{h} \tag{4.14}$$

using the operators $\boldsymbol{E}_r$. The total volume is obtained by summing the contributions from each node of the spin network graph Γ contained in the region R. The polyhedral correspondence described in the last section is recent and previous work considered the region of space surrounding a node to be a cell dual to the spin network graph—dual in a Poincaré sense of having faces transverse to the network links and volume where the graph has a node.

Due to the necessity of a regularization scheme there are different proposals for the volume operator at a node. The operator originally proposed by Rovelli and Smolin [121] is

$$\hat{V}_{\mathrm{RS}} = \alpha \sqrt{\sum_{r<s<t} \left|\boldsymbol{E}_r \cdot (\boldsymbol{E}_s \times \boldsymbol{E}_t)\right|}, \tag{4.15}$$

where $\alpha > 0$ is a multiplicative constant and the sum over $r < s < t$ spans triples of links at the node. A second operator introduced by Ashtekar and Lewandowski [122] is

$$\hat{V}_{\mathrm{AL}} = \alpha \sqrt{\left| \sum_{r<s<t} \varepsilon(v^r, v^s, v^t)\ \boldsymbol{E}_r \cdot (\boldsymbol{E}_s \times \boldsymbol{E}_t)\right|}, \tag{4.16}$$

where v^r are the tangents to the links of the embedded spin network Γ. The ε are signs $\varepsilon(v^r, v^s, v^t) = \pm 1$ corresponding to whether the triple is right or left handed respectively. When $\hat{V}_{\mathrm{AL}}$ is viewed as an operator on the intertwiner space the tangents

v^r and the signs $\varepsilon(v^r, v^s, v^t)$ have to be understood as external fixed data.

Both the Rovelli-Smolin and the Ashtekar-Lewandowski proposals, defined here on the intertwiner space, admit a classical versions. In the reverse of the process described at the end of the last section we dequantize the operators $\boldsymbol{E}_r$ to obtain vectors $\boldsymbol{A}_r$. This results in two distinct functions on phase space:

$$V_{\mathrm{RS}}(\boldsymbol{A}_r) \qquad \text{and} \qquad V_{\mathrm{AL}}(\boldsymbol{A}_r). \tag{4.17}$$

Recently, a third proposal for the volume operator at a node has emerged [109]. Motivated by the geometry of the Minkowski theorem discussed in the last section, Doná, Bianchi and Speziale suggest the promotion of the classical volume of the polyhedron associated to $\{\boldsymbol{A}_r\}$ to an operator

$$V_{\mathrm{Pol}}(\boldsymbol{A}_r) \qquad \rightarrow \qquad \hat{V}_{\mathrm{Pol}}(\boldsymbol{E}_r). \tag{4.18}$$

In the case of a 4-valent node all three of these proposals agree. The heart of this chapter is a study of the semiclassics of this operator. Already in the 5-valent case these operators start to differ and the last part of the chapter initiates the study of the classical function $V_{\mathrm{Pol}}(\boldsymbol{A}_r)$. Now we can also see why 3-valent nodes do not carry volume. Gauge symmetry imposes $\boldsymbol{A}_1 + \boldsymbol{A}_2 + \boldsymbol{A}_3 = 0$ on the classical side and $\boldsymbol{E}_1 + \boldsymbol{E}_2 + \boldsymbol{E}_3 = 0$ in terms of the quantum operators, in either case the triple product defining the volume vanishes.

The three proposals discussed here differ in important respects. Most importantly the Ashtekar-Lewandowski proposal depends on the embedding of the spin network graph Γ into Σ due to its dependencies on the tangent vectos v^r. Until significant reasons for preferring the embedded picture emerge, e.g. it is possible that the knotting of the graph in the embedded picture may be meaningful, this seems like an undesirable feature. Nonetheless, a significant body of research [117, 118, 119, 123] has been amassed in the service of efficiently calculating these signs and this work provides ample data on $\hat{V}_{\mathrm{AL}}$ that will be useful for future comparisons.

In a previous semiclassical study of the volume operator Thiemann and Flori [124] argue that the Rovelli-Smolin and the Ashtekar-Lewandowski proposals share a common defect. Because of the subgraph inclusion discussed in the introduction and on physical grounds one would expect the volume operator of an F-valent node to limit to the volume operator of an $(F-1)$-valent node as you take the spin j_r labeling the rth link to zero. Neither of these proposals has this property, which is called cylindrical consistency. On the other hand, because of its geometrical definition the operator $\hat{V}_{\mathrm{Pol}}$ automatically has this property. However, in the section on 5-valent nodes we will see that this is implemented in a non-trivial fashion. Continued study of the 5-valent case should shed further light on the differences between these operators. We turn now to the 4-valent case.

4.3.1 The volume of a quantum tetrahedron

In the case of a node with four links, $F = 4$, all the proposals for the volume operator discussed above *coincide* and match the operator introduced by Barbieri [125] for the volume of a quantum tetrahedron. The Hilbert space $\mathcal{K}_4$ of a quantum tetrahedron is the intertwiner space of four representations of $SU(2)$,

$$\mathcal{K}_4 = \mathrm{Inv}\left(\mathcal{H}_{j_1} \otimes \mathcal{H}_{j_2} \otimes \mathcal{H}_{j_3} \otimes \mathcal{H}_{j_4}\right). \tag{4.19}$$

We introduce a basis into this Hilbert space using the recoupling channel $\mathcal{H}_{j_1} \otimes \mathcal{H}_{j_2}$ and call these basis states $|k\rangle$. The basis vectors are defined as

$$|k\rangle = \sum_{m_1 \cdots m_4} i_k^{m_1 m_2 m_3 m_4} |j_1, m_1\rangle |j_2, m_2\rangle |j_3, m_3\rangle |j_4, m_4\rangle, \tag{4.20}$$

where the tensor $i_k^{m_1 m_2 m_3 m_4}$ is defined in terms of the Wigner $3j$-symbols as

$$i_k^{m_1 m_2 m_3 m_4} = \sqrt{2k+1} \sum_{m=-k}^{k} (-1)^{k-m} \begin{pmatrix} j_1 & j_2 & k \\ m_1 & m_2 & m \end{pmatrix} \begin{pmatrix} k & j_3 & j_4 \\ -m & m_3 & m_4 \end{pmatrix}. \tag{4.21}$$

The index k ranges from $k_{\min}$ to $k_{\max}$ in integer steps with,

$$k_{\min} = \max(|j_1 - j_2|,\, |j_3 - j_4|) \qquad \text{and} \qquad k_{\max} = \min(j_1 + j_2,\, j_3 + j_4)\,. \tag{4.22}$$

The dimension d of the Hilbert space $\mathcal{K}_4$ is finite and given by

$$d = k_{\max} - k_{\min} + 1.^3 \tag{4.23}$$

The states $|k\rangle$ form an orthonormal basis of eigenstates of the operator $\boldsymbol{E}_r \cdot \boldsymbol{E}_s$. This operator measures the dihedral angle between the faces r and s of the quantum tetrahedron [126].

The operator $\sqrt{\boldsymbol{E}_r \cdot \boldsymbol{E}_r}$ measures the area of the rth face of the quantum tetrahedron and states in $\mathcal{K}_4$ are area eigenstates with eigenvalues $8\pi\gamma L_P^2 \sqrt{j_r(j_r+1)}$,

$$\sqrt{\boldsymbol{E}_a \cdot \boldsymbol{E}_a}\; |i\rangle = 8\pi\gamma L_P^2 \sqrt{j_a(j_a+1)}\; |i\rangle\,. \tag{4.24}$$

The volume operator introduced by Barbieri is

$$\hat{V} = \frac{\sqrt{2}}{3} \sqrt{\left|\boldsymbol{E}_1 \cdot (\boldsymbol{E}_2 \times \boldsymbol{E}_3)\right|}. \tag{4.25}$$

and because of the closure relation

$$(\boldsymbol{E}_1 + \boldsymbol{E}_2 + \boldsymbol{E}_3 + \boldsymbol{E}_4)|i\rangle = 0 \tag{4.26}$$

[3] This dimension can also be expressed in a symmetrical manner that treats all four j_r on an equal footing, $d = \min(2j_1, 2j_2, 2j_3, 2j_4, j_1+j_2+j_3-j_4, j_1+j_2-j_3+j_4, j_1-j_2+j_3+j_4, -j_1+j_2+j_3+j_4)+1$.

this operator coincides with the Rovelli-Smolin operator for $\alpha = 2\sqrt{2}/3$. It also follows from (4.26) that the Ashtekar-Lewandowski operator on $\mathcal{K}_4$ is simply given by $\hat{V}_{AL} = \sqrt{|\sigma|}\hat{V}_{RS}$, where σ is a number depending on the Grot-Rovelli class [127] of the link tangents at the node that can attain the values $\sigma = 0, \pm 1, \pm 2, \pm 3, \pm 4$.[4] Therefore the Ashtekar-Lewandowski volume operator coincides numerically with the operators (4.15) and (4.18) when the tangents to the links fall into the class corresponding to $\sigma = \pm 4$ and the constant α is choosen to be $2\sqrt{2}/3$ as before. Otherwise, it is proportional to it. The volume operator introduce by Barbieri can be understood as a special case of the volume of a quantum polyhedron discussed in [109].

In order to compute the spectrum of the volume operator, it is useful to introduce the operator $\hat{Q}$ defined as

$$\hat{Q} = \frac{2}{9}\, \boldsymbol{E}_1 \cdot (\boldsymbol{E}_2 \times \boldsymbol{E}_3). \tag{4.28}$$

It represents the square of the oriented volume. The matrix elements of this operator are easily computed and we report them momentarily. The eigenstates $|q\rangle$ of the operator $\hat{Q}$,

$$\hat{Q}\,|q\rangle = q\,|q\rangle, \tag{4.29}$$

are also eigenstates of the volume. The eigenvalues of the volume are simply given by the square-root of the modulus of q,

$$\hat{V}\,|q\rangle = \sqrt{|q|}\,|q\rangle. \tag{4.30}$$

The matrix elements of the operator $\hat{Q}$ in the basis $|k\rangle$ were originally computed independently by Chakrabarti and then by Lévy-Leblond and Lévy-Nahas [111, 112]. They are given by

$$\hat{Q} = (8\pi\gamma L_P^2)^3 \sum_{k=k_{\min}+1}^{k_{\max}} 2i\, \frac{\Delta(k, A_1, A_2)\Delta(k, A_3, A_4)}{\sqrt{k^2 - 1/4}} \Big(|k\rangle\langle k-1| - |k-1\rangle\langle k| \Big) \tag{4.31}$$

here we introduce the shorthand $A_r = j_r + 1/2$, indeed in the semiclassical limit the operator $\boldsymbol{J}_r$ with eigenvalue j_r is associated to an angular momentum vector $\boldsymbol{A}_r$ with magnitude $A_r = j_r + 1/2$, see [37] for more detailed discussion of this point. The function $\Delta(a, b, c)$ returns the area of a triangle with sides of length (a, b, c) and is conveniently expressed in terms of Heron's formula

$$\Delta(a, b, c) = \frac{1}{4}\sqrt{(a+b+c)(a+b-c)(a-b+c)(-a+b+c)}. \tag{4.32}$$

[4]The number σ is defined as

$$\sigma(v^1, v^2, v^3, v^4) = \varepsilon(v^1, v^2, v^3) - \varepsilon(v^1, v^2, v^4) + \varepsilon(v^1, v^3, v^4) - \varepsilon(v^2, v^3, v^4). \tag{4.27}$$

Computing the spectrum of $\hat{Q}$ amounts to computing the eigenvalues of a $d \times d$ matrix, where d is the dimension of the Hilbert space given in (4.23). This can be done numerically and several of our figures compare the eigenevalues calculated in this manner to the results of the Bohr-Sommerfeld quantization, see section 4.5.

There are a number of properties of the spectrum of $\hat{Q}$ (and therefore of $\hat{V}$) that can be determined analytically. We list some of them below and refer to Lévy-Leblond and Lévy-Nahas [112] for a detailed analysis:

- The spectrum of $\hat{Q}$ is non-degenerate: it contains d distinct real eigenvalues. This is a consequence of the fact that the matrix elements of $\hat{Q}$ on the basis $|k\rangle$ determine a $d \times d$ Hermitian matrix of the form

$$\begin{pmatrix} 0 & ia_1 & 0 & \cdots \\ -ia_1 & 0 & ia_2 & \ddots \\ 0 & -ia_2 & 0 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix} \tag{4.33}$$

  with real coefficients a_i.

- The non-vanishing eigenvalues of $\hat{Q}$ come in pairs $\pm q$. A vanishing eigenvalue is present only when the dimension d of the intertwiner space is odd. These two properties are also a consequence of the structure of the matrix (4.33) discussed above. It is interesting to derive them in the following way.

  Let us introduce the time reversal operator for angular momenta $\hat{\Theta}$ (this operator was introduced in section 3.4.2). By conjugation, it sends the vectors $\boldsymbol{E}_a$ to $-\boldsymbol{E}_a$, or equivalently a state $|j_a, m_a\rangle$ in $|j_a, -m_a\rangle$. Using the definition (4.20) of the basis $|k\rangle$ of intertwiner space, we find that the time reversal operator is given by

$$\hat{\Theta} = \sum_{k=k_{\min}}^{k_{\max}} (-1)^k \, |k\rangle\langle k| \, . \tag{4.34}$$

  Clearly, the oriented volume operator $\hat{Q}$ is odd under time reversal,

$$\hat{\Theta}\hat{Q}\hat{\Theta} = -\hat{Q} \, . \tag{4.35}$$

  As a consequence, if $|q\rangle$ is an eigenstate of $\hat{Q}$ with eigenvalue q then $\hat{\Theta}|q\rangle$ is an eigenstate with eigenvalue $-q$. The non-degeneracy of the spectrum implies that 0 can appear as an eigenvalue only if the dimension d is odd. As a result, the non-vanishing eigenvalues of the volume operator are twice-degenerate, and a non-degenerate vanishing eigenvalue is present only for odd dimension d.

- For given spins $j_1, \ldots, j_4$, Brunnemann and Thiemann have estimated the maximum volume eigenvalue using Gershgorin's circle theorem [117] and they find that it scales as $v_{\max} \sim j_{\max}^{3/2}$, where $j_{\max}$ is the largest of the four spins j_r.

- They have also estimated the minimum non-vanishing eigenvalue (volume gap) and find that it scales as $v_{\min} \sim j_{\max}^{1/2}$ [117, 119].

Our semiclassical analysis reproduces all of these results and provides several new insights into their structure. This completes our review of the various proposals for the volume operator in loop gravity and we turn now to the classical geometry of tetrahedra.

4.4 Tetrahedral volume on shape space

In this section we discuss, in some detail, the classical analog of formula (4.25), the volume of a tetrahedron as a function on the shape phase space $\mathcal{P}(A_1, \ldots, A_4) \equiv \mathcal{P}_4$. This will be the starting point of our Bohr-Sommerfeld analysis in the next section.

As discussed above the Minkowski theorem guarantees the existence and uniqueness of a tetrahedron associated to any four vectors $\boldsymbol{A}_r$, $(r = 1, \ldots, 4)$ that satisfy $\boldsymbol{A}_1 + \cdots + \boldsymbol{A}_4 = \boldsymbol{0}$. The magnitudes $A_r \equiv |\boldsymbol{A}_r|, (i = 1, \ldots, 4)$ are interpreted as the face areas and the directions $\hat{A}_i, (i = 1, \ldots, 4)$ are unit outward pointing normals to the faces. Without loss of generality we will take $A_1 \leq A_2 \leq A_3 \leq A_4$. In terms of these magnitudes, a condition for the existence of a tetrahedron is that $A_1 + A_2 + A_3 \geq A_4$, equality yielding a flat (zero volume) tetrahedron. This is clearly necessary, as there would be no way to satisfy closure if $A_1 + A_2 + A_3 < A_4$ held. It is not difficult to argue that this is also sufficient for there to exist at least one tetrahedron with these face areas (in fact, there are infinitely many). The space of tetrahedra with four fixed face areas $\mathcal{P}(A_1, A_2, A_3, A_4) \equiv \mathcal{P}_4$ is, as we will now argue, a sphere.

Following the general construction outlined in section 4.2, the canonical coordinates on $\mathcal{P}_4$ are $\mu_1 = |\boldsymbol{A}_1 + \boldsymbol{A}_2|$ and ϕ, the angle between $\boldsymbol{v}_1 = \boldsymbol{A}_1 \times \boldsymbol{A}_2$ and $\boldsymbol{w}_1 = (\boldsymbol{A}_1 + \boldsymbol{A}_2) \times \boldsymbol{A}_3$. For the remainder of the chapter we adopt the simpler notation $\boldsymbol{A} \equiv \boldsymbol{A}_1 + \boldsymbol{A}_2$ and $A = |\boldsymbol{A}_1 + \boldsymbol{A}_2| = \mu_1$. Recalling that the $\boldsymbol{A}_i, (i = 1, \ldots, 4)$ vectors are to be thought of as generators of $SU(2)$ actions we observe that A generates rotations of $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ about the $\hat{A}$-axis. This action rotates $\boldsymbol{v}_1$ about the perpendicular $\hat{A}$-axis while leaving $\boldsymbol{w}_1$ fixed and thus increments the angle ϕ. This is the geometrical content of the poisson bracket relation $\{A, \phi\} = 1$. Because $\boldsymbol{A}$ is fixed by this rotation the closure condition,

$$\boldsymbol{A}_1 + \boldsymbol{A}_2 + \boldsymbol{A}_3 + \boldsymbol{A}_4 = \boldsymbol{A} + \boldsymbol{A}_3 + \boldsymbol{A}_4 = 0 \tag{4.36}$$

is also unaffected by such a rotation. Consequently, we have a whole circle of distinct tetrahedra for each value of A with $A_{\min} \leq A \leq A_{\max}$, where $A_{\min} \equiv \max\{A_2 - A_1, A_4 - A_3\}$ and $A_{\max} \equiv \min\{A_2 + A_1, A_4 + A_3\}$. The collection of these circles over the interval of allowed A values is the slicing of a sphere into lines of latitude over the range of its z diameter. This is made precise by the behavior at the ends

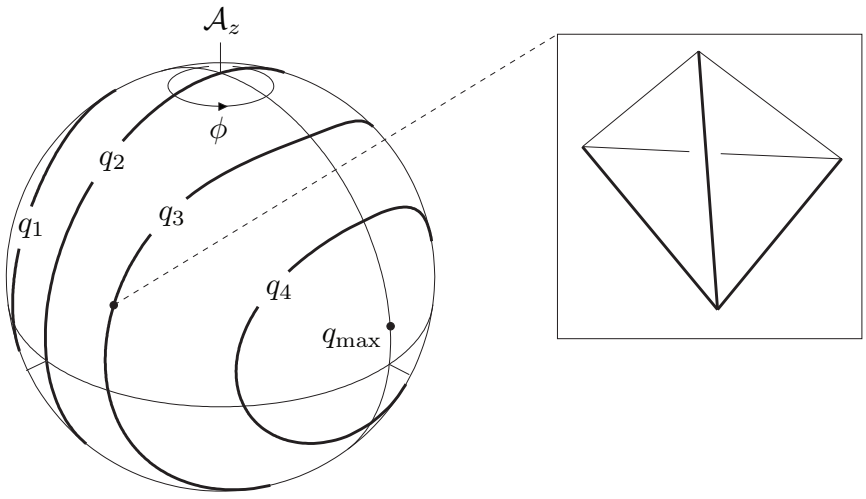


Figure 4.1: The space of shapes for a tetrahedron, $\mathcal{P}(\frac{3}{2},\frac{3}{2},\frac{3}{2},\frac{3}{2})$. The darkened contours are quantized level sets of the classical volume squared, Q. One quantized level set is hidden from view. The inset depicts a tetrahedron corresponding to the dot on the q_3 level set. The point corresponding to the tetrahedron with the largest possible volume, given the face area constraints, is also marked with a dot. The shape space coordinates are indicated at the top of the sphere.

of the range of A; either the vectors $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ or the vectors $\boldsymbol{A}_3$ and $\boldsymbol{A}_4$ become colinear, and consequently the four A-vectors are coplanar. Such configurations are all equivalent up to overall rotations in $\mathbb{R}^3$ and hence under the Kapovich-Millson reduction (section 4.2) they correspond to a single point in the reduced space. As usual, the ϕ coordinate becomes ill defined at these poles, here this is because either $\boldsymbol{v}_1$ or $\boldsymbol{w}_1$ vanishes.

The shape space sphere is depicted in Figure 4.1. The embedding space is $\mathbb{R}^3$ and can be thought of as a copy of the $SU(2)$ Lie algebra associated to $\boldsymbol{A}$, we will call this angular momentum space $\mathcal{A}$. Coordinates on $\mathcal{A}$ and the shape space sphere are defined as follows: place the origin of $\mathcal{A}$ at the center of the range of A and choose cartesian coordinates $\mathcal{A}_x$, $\mathcal{A}_y$ and $\mathcal{A}_z$ with $\mathcal{A}_z = A - (A_{\max} + A_{\min})/2$. The radius of the shape space sphere is $r = (A_{\max} - A_{\min})/2$ and the relation $\mathcal{A}_z = r\cos\theta$ together with the angle ϕ from above define a spherical coordinate system on the sphere. The choice of coordinates settled we proceed to the volume operator on the shape space.

As discussed in section 4.3 we will be considering the classical volume,

$$V = \frac{\sqrt{2}}{3}\sqrt{|\boldsymbol{A}_1 \cdot (\boldsymbol{A}_2 \times \boldsymbol{A}_3)|}, \tag{4.37}$$

and just as with the quantum theory it will be more straightforward to work with the

squared classical volume,

$$Q = \frac{2}{9}\boldsymbol{A}_1 \cdot (\boldsymbol{A}_2 \times \boldsymbol{A}_3). \tag{4.38}$$

Because Q is a rotational invariant it projects onto the shape phase space and can be thought of as a function of the A and ϕ coordinates, $Q(A, \phi)$.[5] This expression is easily derived by computing $\boldsymbol{v}_1 \times \boldsymbol{w}_1$; from the definitions of $\boldsymbol{v}_1$ and $\boldsymbol{w}_1$ one finds

$$\boldsymbol{v}_1 \times \boldsymbol{w}_1 = \frac{9}{2} Q \boldsymbol{A}. \tag{4.39}$$

Note that the magnitude $|\boldsymbol{v}_1| = |\boldsymbol{A}_1 \times \boldsymbol{A}_2|$ is equal to twice the area Δ of a triangle with side lengths A_1, A_2 and A and, using the closure relation (4.36), similarly $|\boldsymbol{w}_1| = |\boldsymbol{A}_3 \times \boldsymbol{A}_4|$ is twice the area $\bar{\Delta}$ of a triangle with side lengths A_3, A_4 and A. The definition of ϕ as the angle between $\boldsymbol{v}_1$ and $\boldsymbol{w}_1$ allows us to conclude that the volume squared is

$$Q = \frac{8}{9}\frac{\Delta\bar{\Delta}}{A}\sin\phi. \tag{4.40}$$

Calculating the areas Δ and $\bar{\Delta}$ using Heron's formula (4.32),

$$\Delta = \frac{1}{4}\sqrt{[(A_1 + A_2)^2 - A^2][A^2 - (A_1 - A_2)^2]}, \tag{4.41}$$

$$\bar{\Delta} = \frac{1}{4}\sqrt{[(A_3 + A_4)^2 - A^2][A^2 - (A_3 - A_4)^2]}, \tag{4.42}$$

shows that for fixed $A_1, \ldots, A_4$, Q is indeed only a function of the coordinates A and ϕ. The expression (4.40) will be the central tool of our Bohr-Sommerfeld quantization. In anticipation of the results of the next section, some quantized level sets of Q are depicted in Figure 4.1.

Before proceeding to the Bohr-Sommerfeld quantization we investigate the extrema of $|Q|$. At the classical level, the minimum of $|Q|$ is always zero: Open the angle ϕ until the plane spanned by $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ coincides with the plane spanned by $\boldsymbol{A}_3$ and $\boldsymbol{A}_4$, then $Q = 0$. This, however, does not lead to the conclusion that the minimum of V is always zero. This is because the Minkowki theorem does not hold for planar configurations of the $\boldsymbol{A}_r$, it is generically singular for these configurations. Certainly $Q = 0$ implies that $V = 0$ the issue is the correspondence between a planar set of $\boldsymbol{A}_r$ and such a flat tetrahedron. Geometrically this is clear, a flat tetrahedron has *faces* that lie in a plane and the normals to these faces are all collinear. Thus it is only the subset of planar configurations of vectors $\boldsymbol{A}_r$ that are actually collinear that can have a Minkowki type correspondence with a flat tetrahedron. We can say more: the collinear vectors must satisfy closure and so, for some choice of signs we must have $\pm A_1 \pm A_2 \pm A_3 \pm A_4 = 0$. Taking into account the ordering convention $A_1 \leq A_2 \leq A_3 \leq A_4$ we can bring the number of cases down to just three, either

[5] Because A and $\mathcal{A}_z$ differ by a constant shift we will freely switch between them.

$A_2 - A_1 = A_4 - A_3$ or $A_2 + A_1 = A_4 + A_3$ or $A_2 + A_1 = A_4 - A_3$. The last condition leads to a trivial shape space consisting of a single point because $A_{\min} = A_{\max}$. We will call these three conditions the "flatness" conditions. They will also be significant in the Bohr-Sommerfeld quantization of the next section.

Note that even for collinear configurations of the $\boldsymbol{A}_r$ the Minkowski theorem still doesn't hold. The trouble is uniqueness. An infinite number of flat configurations all share the same area vectors. In fact, the differential structure of the shape space breaks down when the flatness conditions are satisfied (for an analogous observation see [22]). The precise treatment of flat configurations warrants more investigation. Notice that everything that has been said up to this point is in regards to a tetrahedron that is exactly flat. This is significant because it highlights the singular nature of the Minkowski construction for planar configurations of the $\boldsymbol{A}_r$. You can construct a tetrahedron with arbitrarily small volume from a set of vectors that is arbitrarily close to planar, you only run into trouble when you ask for exactly flat tetrahedra. For the purposes of the present work we summarize the preceding observations: unless the flatness conditions are satisfied, $Q = 0$ should not lead to the conclusion that there is a constructible tetrahedron with $V = 0$; if one or two of the flatness conditions are satisfied then only the corresponding pole of the shape space sphere corresponds to flat configurations with $V = 0$ and they generally correspond to a whole class of such tetrahedra; thus the great circle on which $Q = 0$ ($\phi \in \{0, \pi\}$) will be regarded as mathematically useful but largely physically meaningless.

Turning to the maxima of Q, we begin with the case where three faces have fixed areas, say A_1, A_2 and A_3 but the full vectors $\boldsymbol{A}_s, (s = 1, 2, 3)$ are not given. Writing the triple product of Q as the determinant of a matrix $M = (\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3)$ whose columns are the vectors $\boldsymbol{A}_1, \boldsymbol{A}_2$ and $\boldsymbol{A}_3$ and squaring yields,

$$Q^2 = \frac{4}{81} \det M^T \det M = \frac{4}{81} \det \begin{pmatrix} A_1^2 & \boldsymbol{A}_1 \cdot \boldsymbol{A}_2 & \boldsymbol{A}_1 \cdot \boldsymbol{A}_3 \\ \boldsymbol{A}_2 \cdot \boldsymbol{A}_1 & A_2^2 & \boldsymbol{A}_2 \cdot \boldsymbol{A}_3 \\ \boldsymbol{A}_3 \cdot \boldsymbol{A}_1 & \boldsymbol{A}_3 \cdot \boldsymbol{A}_2 & A_3^2 \end{pmatrix}, \tag{4.43}$$

where M^T denotes the transpose of M. Taking the unknown dot products $\boldsymbol{A}_s \cdot \boldsymbol{A}_t, (s < t = 1, 2, 3)$ as variables and extremizing one finds the minima already discussed, where $\boldsymbol{A}_1$, $\boldsymbol{A}_2$ and $\boldsymbol{A}_3$ are collinear, and a single global maximum where $\boldsymbol{A}_s \cdot \boldsymbol{A}_t = 0, (s < t = 1, 2, 3)$. This maximum must satisfy closure and so,

$$A_4^2 = A_1^2 + A_2^2 + A_3^2 + 2\boldsymbol{A}_1 \cdot \boldsymbol{A}_2 + 2\boldsymbol{A}_1 \cdot \boldsymbol{A}_3 + 2\boldsymbol{A}_2 \cdot \boldsymbol{A}_3 = A_1^2 + A_2^2 + A_3^2. \tag{4.44}$$

Then the maximum volume of a tetrahedron with three fixed face areas is the one with three right dihedral angles and the fourth face area given by the equation above. If we use the same technique to maximize the volume over the space where all four face areas are given then the closure condition must be implemented as a constraint. For fixed A_1, A_2 and A_3 it is clear that the constrained maximum will not be larger than the one just found and will be equal to it when the fixed value of A_4 is that

of (4.44). Rather than implementing the constraint with a lagrange multiplier it is easier to extremize equation (4.40), which treats the $A_r, (r = 1, \cdots 4)$ on an equal footing. Introduce the shorthand $P_0(A^2)$ for the polynomial part of the area product $\Delta\bar{\Delta}$,

$$P_0(A^2) \equiv [A^2-(A_1-A_2)^2][A^2-(A_3-A_4)^2][(A_1+A_2)^2-A^2][(A_3+A_4)^2-A^2], \quad (4.45)$$

that is, $\Delta\bar{\Delta} = 1/16\sqrt{P_0(A^2)}$. This is conveniently viewed as a quartic polynomial in $x \equiv A^2$ with roots $\bar{r}_1 \equiv (A_1 - A_2)^2$, $\bar{r}_2 \equiv (A_3 - A_4)^2$, $\bar{r}_3 \equiv (A_1 + A_2)^2$ and $\bar{r}_4 \equiv (A_3 + A_4)^2$. The expression for the squared volume (4.40) simplifies to $Q = 1/(18\sqrt{x})\sqrt{P_0(x)}\sin\phi$. This expression is maximized when $\phi = \pi/2$ and when $\partial Q/\partial A = 0$ or,

$$P_0(x) = xP_0'(x). \quad (4.46)$$

This condition is another quartic equation, the roots of which are not worth explicitly displaying in general, however choosing the root which maximizes Q, say $\bar{x}$, then we have $Q_{\max} = 1/18\sqrt{P_0'(\bar{x})}$, an expression that will be useful below. This concludes our general treatment of the extrema of Q. It would be interesting to investigate the extrema of Q imposing closure but without fixing any of the face areas, we leave this for future work.

4.5 Bohr-Sommerfeld quantization of tetrahedra

As discussed in Chapter 2, the later development of Bohr's intuitive correspondence principle by Sommerfeld and Ehrenfest lead to an elegant approximate quantization, now called Bohr-Sommerfeld quantization. Because the calculations of this section are quite detailed we sketch the approach. Our first order of business is to find the orbits of the dynamics generated by Q, several of these orbits have already been displayed in Figure 4.1. The Bohr-Sommerfeld quantization condition is expressed in terms of the action I associated to the each of these orbits:

$$I(q) = \oint A d\phi = 2\pi n, \quad (4.47)$$

here we have denoted the level value of Q by q, the quantization level by n and expressed the action in terms of the shape space coordinates (for the details see below). We also take units in which $\hbar = 1$ and so $h = 2\pi$. In the main body of the section we calculate I and impose this quantization condition. This is the usual Bohr-Sommerfeld procedure with Q playing the role of the Hamiltonian H and q the role of the energy E.

The classical evolution of the area tetrahedron with Q taken to be the Hamiltonian is easiest to calculate in terms of A^2,

$$\frac{d(A^2)}{d\lambda} \equiv \{A^2, Q\} = 2A\{A, Q\} = 2A\frac{\partial Q}{\partial \phi}, \quad (4.48)$$

where λ is defined to be the variable conjugate to Q and ϕ is, as in the previous section, the angle conjugate to A. The right hand side can be evaluated by differentiating (4.40) with respect to ϕ,

$$\frac{d(A^2)}{d\lambda} = \frac{16}{9}\Delta\bar{\Delta}\cos\phi. \tag{4.49}$$

Equation (4.40) can be used again to eliminate the cosine function,

$$\frac{d(A^2)}{d\lambda} = \frac{1}{9}\sqrt{(4\Delta)^2(4\bar{\Delta})^2 - (2A)^2(9Q)^2}. \tag{4.50}$$

The argument of the square root plays an important role in what follows and so we introduce a shorthand for it,

$$P(A^2, Q^2) \equiv (4\Delta)^2(4\bar{\Delta})^2 - (2A)^2(9Q)^2. \tag{4.51}$$

This is a quartic polynomial in A^2 and can be expressed in terms of $P_0(A^2)$,

$$P(A^2, Q^2) = P_0(A^2) - (2A)^2(9Q)^2, \tag{4.52}$$

which was defined at (4.45):

$$P_0(A^2) \equiv [A^2-(A_1-A_2)^2][A^2-(A_3-A_4)^2][(A_1+A_2)^2-A^2][(A_3+A_4)^2-A^2]. \tag{4.53}$$

Taking the short hand $x \equiv A^2$ we can separate variables in (4.50) and integrate to find,

$$\lambda(x) = 9\int_{r_2}^{x} \frac{d\tilde{x}}{\sqrt{(\tilde{x}-r_1)(\tilde{x}-r_2)(r_3-\tilde{x})(r_4-\tilde{x})}}, \tag{4.54}$$

we assume that the four distinct real roots (r_1, r_2, r_3, r_4) of the quartic $P(x \equiv A^2, Q^2)$ are ordered as $r_1 < r_2 < r_3 < r_4$. This is an elliptic integral; to bring it to the standard Jacobi form we use a Möbius transformation that brings the quartic to a conventional one with roots $\pm 1, \pm\frac{1}{\sqrt{m}}$.[6] This is possible as long as the cross-ratio of the r's is the same as the cross-ratio of the conventional roots, we use this to set the elliptic parameter m. Explicitly the substitution is

$$z^2 = \frac{(r_4-r_2)(r_3-x)}{(r_3-r_2)(r_4-x)} \qquad \text{and} \qquad m = \frac{(r_3-r_2)(r_4-r_1)}{(r_4-r_2)(r_3-r_1)}. \tag{4.55}$$

After evaluation of the integral and some algebra this leads to the solution,

$$x(\lambda) = A^2(\lambda) = \frac{r_3(r_4-r_2) - r_4(r_3-r_2)\mathrm{sn}^2\left(\frac{\lambda}{9g}, m\right)}{(r_4-r_2) - (r_3-r_2)\mathrm{sn}^2\left(\frac{\lambda}{9g}, m\right)}, \tag{4.56}$$

[6] To avoid notational conflicts with the intertwiner eigenstates $|k\rangle$ all elliptic functions are written in terms of the elliptic parameter $m \equiv k^2$ instead of the elliptic modulus k.

with

$$g = \frac{2}{\sqrt{(r_4 - r_2)(r_3 - r_1)}}. \tag{4.57}$$

This is a complete solution of the dynamics. After the specification of a value for the volume, the quartic $P(x, Q^2)$ can be solved and the volume evolution of the intermediate coupling A is given by (4.56). The evolution is periodic and the period can be expressed in terms of the complete elliptic integral of the first kind $K(m)$ by $T = 9g \times 2K = 18gK$. The fundamental period of the elliptic functions, $T_0 = 4K$, is halved because they appear squared. For definiteness, in what follows we will assume that the elliptic parameter m is less than one. If this is not the case apply the transformation $\mathrm{sn}\,(u, m) = \frac{1}{\sqrt{m}}\mathrm{sn}\,(\sqrt{m}u, \frac{1}{\sqrt{m}})$ and you will find that the effect on (4.56) and (4.57) is to switch the roles of r_1 and r_2 throughout.

Before proceeding to the calculation of the action I of a curve $A(\lambda)$, we pause to describe some of the properties of the quartic $P(x, Q^2)$ that will be useful in this calculation. For the value $Q^2 = 0$ the quartic simplifies and is given by $P_0(x)$. In particular $P_0(x)$ can be explicitly factored, see (4.53), and we will call its four positive real roots $\bar{r}_1 < \bar{r}_2 < \bar{r}_3 < \bar{r}_4$. For non-zero real values of the volume Q, the roots of the quartic equation $P(x, Q^2) = 0$ are given by the intersections of the line $y = 324Q^2x$ with the quartic $y = P(x, 0)$. This allows us to say a few general things about the roots $\{r_1, r_2, r_3, r_4\}$ of $P(x, Q^2)$. The largest and smallest roots are always real and satisfy $0 < r_1 < \bar{r}_1 = \min\{(A_1 - A_2)^2, (A_3 - A_4)^2\}$ and $r_4 > \bar{r}_4 = \max\{(A_1 + A_2)^2, (A_3 + A_4)^2\}$. Meanwhile for small enough Q the middle two roots are also real and satisfy, $r_2 > \bar{r}_2 = \max\{(A_1 - A_2)^2, (A_3 - A_4)^2\}$ and $r_3 < \bar{r}_3 = \min\{(A_1 + A_2)^2, (A_3 + A_4)^2\}$. As Q grows the middle two roots coalesce and then go off into the complex plane. These observations are summarized graphically in the first two panels of Figure 4.5.

The roots of a polynomial coalesce when the polynomial and its first derivative simultaneously vanish, $P(x) = P'(x) = 0$. With the notation introduced above $P(x) \equiv P_0(x) - 324xQ^2$ (we suppress the Q dependence), two roots will coalesce when $P(x) = 0$ and, $dP/dx = 0 = P_0'(x) - 324Q^2$ or $Q_{\mathrm{coal}} = 1/18\sqrt{P_0'(x)}$ both hold. The latter is precisely the same condition that we found below equation (4.46) for the maximum volume of a tetrahedron with four fixed face areas,

$$Q_{\max} = 1/18\sqrt{P_0'(x)}. \tag{4.58}$$

This means that the quartic roots coalesce precisely when the maximum real volume of the tetrahedron is achieved and so for real volumes we need only consider real positive roots. We will find that the action is given by complete elliptic integrals and it will be useful to be able to assume that the roots that arise in these formulas are real and positive.

One final observation before we proceed with the Bohr-Sommerfeld quantization of the volume operator, this one nonessential. Because of the simple spherical topology

of the shape space the action $I = \oint A d\phi$ of an orbit γ can be interpreted as the symplectic area contained within this orbit. As usual this is a simple consequence of stokes theorem. The symplectic form on shape space is determined by the poisson bracket relation $\{A, \phi\} = 1$ and is $\omega = dA \wedge d\phi$ so that

$$I = \oint A(\phi) d\phi = \int \omega = \int dA \wedge d\phi. \tag{4.59}$$

We can just as well work with $d\mathcal{A}_z \wedge d\phi = r d\cos\theta \wedge d\phi$ because A and $\mathcal{A}_z$ differ by a constant and so this symplectic area only differs from the solid angle on the sphere by a normalization factor r.

We turn now to the calculation of the action. The action integral for an orbit γ can be re-expressed in terms of conjugate variable λ of the volume,

$$I = \oint A(\phi) d\phi = \oint A(\lambda) \frac{d\phi}{d\lambda} d\lambda. \tag{4.60}$$

Once again turning to (4.40) and solving for ϕ we have, $\phi = \arcsin\left(9AQ/(8\Delta\bar{\Delta})\right)$ and differentiating with respect to λ yields,

$$\begin{aligned} \frac{d\phi}{d\lambda} &= \frac{1}{\sqrt{1 - (9AQ/8\Delta\bar{\Delta})^2}} \left(\frac{9Q}{8\Delta\bar{\Delta}} \frac{dA}{d\lambda} - \frac{9AQ}{8(\Delta\bar{\Delta})^2} \frac{d(\Delta\bar{\Delta})}{d\lambda} \right) \\ &= \left(\frac{Q}{A} - \frac{Q}{\Delta\bar{\Delta}} \frac{d(\Delta\bar{\Delta})}{dA} \right). \end{aligned} \tag{4.61}$$

Returning to the expressions for Δ and $\bar{\Delta}$ ((4.41) and (4.42)) one can calculate $d(\Delta\bar{\Delta})/dA$ and obtain,

$$\begin{aligned} \oint A \frac{d\phi}{d\lambda} d\lambda = \oint Q d\lambda - \oint Q \Bigg(&\frac{A^2}{(A^2 - (A_1 + A_2)^2)} + \frac{A^2}{(A^2 - (A_1 - A_2)^2)} \\ &+ \frac{A^2}{(A^2 - (A_3 + A_4)^2)} + \frac{A^2}{(A^2 - (A_3 - A_4)^2)} \Bigg) d\lambda. \end{aligned} \tag{4.62}$$

Because Q is a constant along the curve the first integral simply gives the period of the elliptic function,

$$\oint Q d\lambda = 18 g Q K(m), \tag{4.63}$$

where again the elliptic parameter m is given by (4.55). The remaining integrals are all of the same type, we introduce a parameter $\bar{r}_i$ which takes the values of the roots at zero volume $(A_1 - A_2)^2, (A_3 - A_4)^2, (A_1 + A_2)^2, (A_3 + A_4)^2$ for $i = 1, \ldots, 4$, and find,

$$\oint Q \frac{A^2}{(A^2 - \bar{r}_i)} d\lambda = Q \oint \frac{r_3(r_4 - r_2) - r_4(r_3 - r_2)\mathrm{sn}^2\left(\frac{\lambda}{9g}, m\right)}{(r_3 - \bar{r}_i)(r_4 - r_2) - (r_4 - \bar{r}_i)(r_3 - r_2)\mathrm{sn}^2\left(\frac{\lambda}{9g}, m\right)} d\lambda \tag{4.64}$$

These integrals are more simply expressed in terms of $u \equiv \lambda/9g$ and the integration over a complete period is over the interval $u \in [0, 2K]$. We find,

$$\oint Q\frac{A^2}{(A^2-\bar{r}_i)}d\lambda = Q\frac{18gr_3}{(r_3-\bar{r}_i)}\int_0^K \frac{1}{1-\alpha_i^2\mathrm{sn}^2(u,m)}du - Q\frac{18gr_4\alpha_i^2}{(r_4-\bar{r}_i)}\int_0^K \frac{\mathrm{sn}^2(u,m)}{1-\alpha_i^2\mathrm{sn}^2(u,m)}du, \tag{4.65}$$

where

$$\alpha_i^2 = \frac{(r_4-\bar{r}_i)(r_3-r_2)}{(r_3-\bar{r}_i)(r_4-r_2)}. \tag{4.66}$$

Collecting all four of these integrals we have,

$$I = 18gQ\left(K - \sum_{i=1}^{4}\left\{\frac{r_3}{(r_3-\bar{r}_i)}\int_0^K \frac{1}{1-\alpha_i^2\mathrm{sn}^2(u,m)}du - \frac{r_4\alpha_i^2}{(r_4-\bar{r}_i)}\int_0^K \frac{\mathrm{sn}^2(u,m)}{1-\alpha_i^2\mathrm{sn}^2(u,m)}du\right\}\right), \tag{4.67}$$

which can be evaluated in terms of complete elliptic integrals yielding,

$$I = 18gQ\left(\left[1-\sum_{i=1}^{4}\frac{r_4}{(r_4-\bar{r}_i)}\right]K(m) - \sum_{i=1}^{4}\frac{\bar{r}_i(r_4-r_3)}{(r_4-\bar{r}_i)(r_3-\bar{r}_i)}\Pi(\alpha_i^2,m)\right), \tag{4.68}$$

where $\Pi(\alpha_i^2, m)$ is the complete elliptic integral of the third kind. To highlight the structure of this result we condense the dependencies on the roots into two coefficients,

$$a \equiv 18g\left[1-\sum_{i=1}^{4}\frac{r_4}{(r_4-\bar{r}_i)}\right] \quad \text{and} \quad b_i \equiv \frac{18g\bar{r}_i(r_4-r_3)}{(r_4-\bar{r}_i)(r_3-\bar{r}_i)} \quad (i=1,\ldots,4). \tag{4.69}$$

This allows us to write I in the more compact form,

$$I = \left(aK(m) - \sum_{i=1}^{4} b_i\Pi(\alpha_i^2,m)\right)Q. \tag{4.70}$$

Figure 4.2 displays a plot of this function for the same parameters used in Figure 4.1.

Despite the action having a closed analytic form, finding the Bohr-Sommerfeld spectrum requires a numerical inversion. Recall that the strategy is to find the volumes for which the corresponding orbits capture $(n + 1/2)2\pi$ worth of area on the sphere. The analytic expression (4.67) has a complicated dependence on the volume; it appears explicitly as an overall multiplicative factor but also implicitly through the elliptic function's dependence on the roots $r_1, \ldots, r_4$, all of which depend on Q.

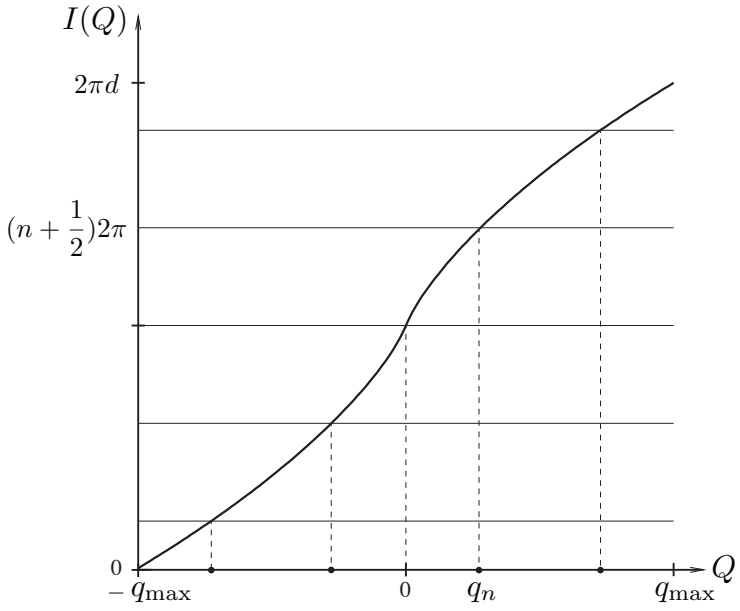


Figure 4.2: A plot of the action integral $I(Q)$ for the same parameters as Figure 4.1. The quantized levels q_n shown satisfy the Bohr-Sommerfeld quantization condition. The corresponding orbits are shown in Figure 4.1.

Instead of trying to invert the action analytically we have calculated its value for several hundred points in the range of classically allowed volumes, interpolated between these values and numerically found the volumes for which the Bohr-Sommerfeld condition is satisfied, this process is illustrated schematically in Figure 4.2. The results of this analysis are presented for two examples in Figure 4.3 along with the numerical diagonalization of the volume matrix elements discussed in section 4.3.1. Recall that $Q = V^2$ and that the eigenvalues satisfy the same relationship $q = v^2$ (see (4.30)).

Section 4.6 contains additional comparisons of the numerical and Bohr-Sommerfeld results. The reduced quality of the approximation for the case where all the spins j_r are equal, exhibited in the lower plot of Figure 4.3 at $j = 4$, is also discussed there.

4.6 Limiting cases for the spectrum

4.6.1 Largest eigenvalues

The limiting behavior of the volume spectrum for large and small eigenvalues can now be explored with the assistance of the analytical formula (4.70). Let us first consider large eigenvalues: the volume function attains a maximum on the sphere and so our strategy will be to expand the action function around this maximum. We

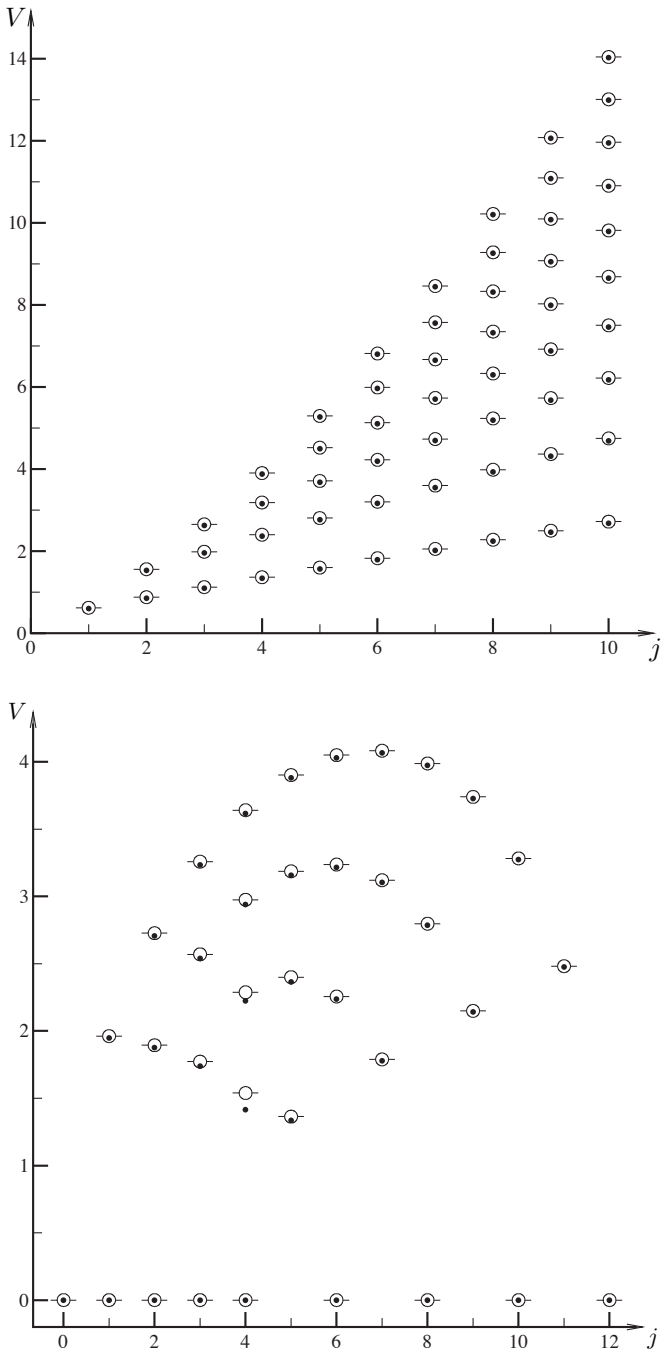


Figure 4.3: A comparison of the Bohr-Sommerfeld and loop gravity Volume spectra. In the upper plot the spins are $\{j, j, j, j+1\}$. In the lower plot the spins are $\{4, 4, 4, j\}$ and j varies in its allowed range. The Bohr-Sommerfeld values for the volume of a tetrahedron are represented by *dots* and the eigenvalues of the loop-gravity volume operator by *circles*. Recall that the spins and areas are related by $A_r = j_r + 1/2$.

have,

$$S(q) = S(q_{\max}) + \left.\frac{\partial S}{\partial q}\right|_{q=q_{\max}} (q - q_{\max}) + \cdots . \tag{4.71}$$

From the theory of action-angle variables, the derivative of the action with respect to the Hamiltonian (in this case the volume) is the period T of the system. For the largest eigenvalue, we are in the same situation as if we were finding a ground state, that is, instead of capturing 2π worth of area on the sphere this state only captures an area π and so,

$$(q_{\max} - q) = \frac{S(q_{\max}) - S(q)}{T(q_{\max})} = \frac{\pi}{T(q_{\max})}. \tag{4.72}$$

Then the largest eigenvalue is given by,

$$q = (q_{\max} - \frac{\pi}{T(q_{\max})}). \tag{4.73}$$

At equation (4.46) we found that the maximum classical volume is attained when,

$$P_0(x) = xP_0'(x) \tag{4.74}$$

or more explicitly when,

$$\frac{1}{x} = \frac{1}{x - (A_1 - A_2)^2} + \frac{1}{x - (A_3 - A_4)^2} + \frac{1}{x - (A_1 + A_2)^2} + \frac{1}{x - (A_3 + A_4)^2}. \tag{4.75}$$

The roots of this quartic for generic A_r are complicated functions of the A_r, however in the case $A_1 = \cdots = A_4 = A_0$ this equation is easily solved and one finds, $x = \frac{4}{3}A_0^2$. The maximum classical volume in this case is $q_{\max} = (2^3/3^{7/2})A_0^3$ and the period is $T(q_{\max}) = 18gK(0) = 18(\sqrt{3}/(2A_0^2))(\pi/2) = 3^{5/2}\pi/(2A_0^2)$. The maximum eigenvalue is given by putting these values into (4.73),

$$v = q^{1/2} = \frac{2^{3/2}}{3^{7/4}} A_0^{3/2} \sqrt{1 - \frac{3}{4A_0}}. \tag{4.76}$$

This reproduces the $A_0^{3/2}$ scaling that has been found in previous works and refines it to the next order. This scaling is plotted as the uppermost line in Figure 4.7. Further corrections could be developed by retaining more terms in (4.71).

Occasionally the equal area tetrahedron has been supposed to be the one whose volume grows most rapidly as the areas are increased. However, it turns out that this is not the case. As discussed in Section 4.4 the tetrahedron with maximum volume depends on the space under consideration. If we consider the space with only three face areas fixed it was found that this is the tetrahedron with three right dihedral angles, $A_1 = A_2 = A_3 = A_0$ and $A_4 = \sqrt{3}A_0$. Indeed these corner tetrahedra have a more rapid scaling with A_0. In Figure 4.4 the Bohr-Sommerfeld spectrum is compared to the numerically calculated exact spectrum and the scaling derived here.

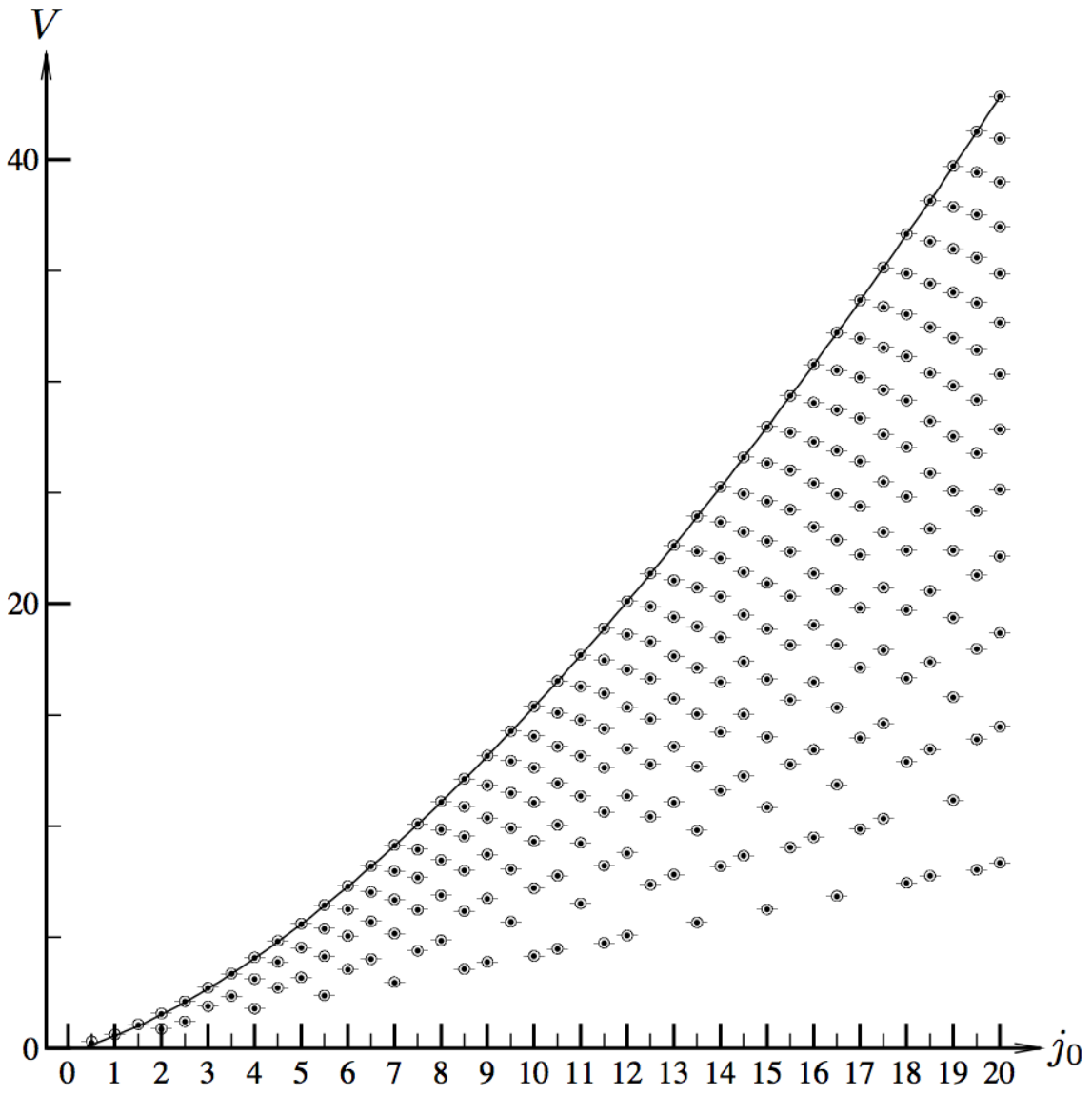


Figure 4.4: This figure compares the semiclassical scaling of the largest volume eigenvalue (dark line) to the Bohr-Sommerfeld (dots) and loop gravity (circles) spectra. The first three spins are given by $j_1 = j_2 = j_3 = j$ and and the largest spin j_4 is given by the closest integer or half integer to $\sqrt{3}j$, that respects the Clebsch-Gordan conditions.

4.6.2 Smallest eigenvalues

The small eigenvalue case has subtleties associated with it. As is clear from the Taylor expansion,

$$S(q) = S(0) - \left.\frac{\partial S}{\partial q}\right|_{q=0} q + \cdots , \tag{4.77}$$

the smallest eigenvalues are associated with the longest period at zero volume, $q_{\min} = (S(0) - S(q))/T(0)$. This would be the end of the story except for the fact that there are a number of shape spaces for which the period at zero volume can become infinitely long and the Taylor expansion above is invalidated. For this reason, we have to treat different shape spaces differently. We will begin by treating the case where the period is finite as it is simpler.

The first subtlety of the small volume cases is immediate: the smallest eigenvalue depends on the dimension of the intertwiner space, $d \equiv \dim \mathrm{Inv}(j_1, \ldots, j_4)$. As already noted, the volume is odd under parity and so there is only a zero eigenvalue when d is odd; this is the only volume state invariant under parity. We're interested in the first non-zero eigenvalue and thus when d is odd the spacing $S(0) - S(q)$ is the spacing between two quantized orbits, h, or in our units 2π. For the finite period case that we are considering the Taylor expansion (4.77) is valid and we have, $T(0)q_{min} = 2\pi$. On the other hand if the phase space is even dimensional, neighboring quantized orbits evenly straddle the zero volume contour and we have $S(0) - S(q) = \pi$, so that $T(0)q_{min} = \pi$.

We found the period above, in general it is $T = 18gK(m)$, where g and m are given in equations (4.55) and (4.56). For $q = 0$ the quartic $P(x, Q = 0)$ factorizes and the roots are $\bar{r}_1 = (A_1 - A_2)^2, \bar{r}_2 = (A_3 - A_4)^2, \bar{r}_3 = (A_1 + A_2)^2, \bar{r}_4 = (A_3 + A_4)^2$, for example. Of course, depending on the choice of $A_1, \ldots, A_4$ other orderings are possible (recall that the r are defined such that $r_1 < r_2 < r_3 < r_4$). If, as suggested after equation (4.56), m is always chosen such that $m < 1$ these other orderings lead to the same result: $g = 1/(2\sqrt{A_1A_2A_3A_4})$ and

$$\begin{aligned} m = &\frac{(A_1 + A_2 + A_3 - A_4)(A_1 + A_2 - A_3 + A_4)}{(2A_1)(2A_2)} \\ &\times \frac{(A_1 - A_2 + A_3 + A_4)(-A_1 + A_2 + A_3 + A_4)}{(2A_3)(2A_4)}. \end{aligned} \tag{4.78}$$

The complete elliptic integral of the first kind has the power series expansion,

$$K(m) = \frac{\pi}{2} \sum_0^\infty \left[\frac{(2n)!}{2^{2n} n!^2}\right]^2 m^n, \tag{4.79}$$

valid for $m < 1$. If m is small enough such that higher order terms can reasonably be neglected then, for d odd,

$$q_{min} \approx \frac{2\pi}{9\pi/2\sqrt{A_1A_2A_3A_4}} = \frac{4}{9}\sqrt{A_1A_2A_3A_4}, \tag{4.80}$$

and this can be improved as much as desired by including more terms from (4.79). Expressed in terms of the volume of the tetrahedron we have,

$$v_{min} \approx (A_1 A_2 A_3 A_4)^{1/4} \begin{cases} 2/3 & \text{if } d \text{ is odd} \\ \sqrt{2}/3 & \text{if } d \text{ is even,} \end{cases} \qquad \text{for } m \ll 1. \tag{4.81}$$

The exact volume eigenvalues derived by Brunnemann and Thiemann [117] are for special cases of the $A_1, \cdots, A_4$, such as $A_1 = A_2 = 1$ and $A_3 = A_4 = j + 1/2$, where only two of the four vectors grow as you increase j. For cases like these we find the same qualitative scaling from this formula $v \sim j^{1/2}$. However, you will notice that this special choice of the A_i leads to $m = 1$ and invalidates the expansion of the elliptic integral $K(m)$, in particular K logarithmically diverges (hence also the period) and a different approach to estimating the eigenvalues is necessary. Brunnemann and Thiemann were lead to consider these special cases by their numerics, they found that these were the phase spaces which lead to the smallest overall values for the volume spectrum. Below we describe what is special about the geometry of these cases and develop an alternative technique for estimating the spectra of these spaces.

The longest periods are achieved when $m = 1$ and the elliptic function theory limits to elementary functions, for example the standard roots of the Jacobi form of the elliptic functions, ± 1 and $\pm 1/\sqrt{m}$, degenerate. We worked out the elliptic modulus in the zero volume limit above (4.78), setting this expression equal to 1 we find four roots, which can be regarded as expressing any one of the A_i in terms of the other three, these are: $A_1 - A_2 = A_3 - A_4$, $A_1 - A_2 = A_4 - A_3$, $A_1 + A_2 = A_3 + A_4$ and $A_1 + A_2 = -A_3 - A_4$, the last of which is clearly unphysical as $A_i \geq 0$. These values have a nice geometrical interpretation in terms of the polynomial $P_0(x)$; they indicate that two of its roots are coalescing. Put this together with our observation that the maximum volume is achieved when $r_2 = r_3$ and consequently when $m = 0$ and we find that $P_0(x)$ is quite useful for characterizing the shape space of a tetrahedron with fixed face areas. These findings are summarized in Figure 4.5.

In the case where $m = 1$ we can no longer use the Taylor expansion and period to find the eigenvalues, instead we have to find the small volume behavior of the action function and try to invert it. The action depends on the volume in two ways, an explicit dependence through the prefactor Q in (4.67), and an implicit dependence through the roots of the quartic $P(x, Q)$. Thus the first step in finding the small volume behavior is to expand the roots as a power series in Q. This would be quite laborious if we had to go through the solutions to the quartic equation, happily a simple alternative exists because we know the roots $\bar{r}_i$ at $Q = 0$. We simply plug $\bar{r}_i + \rho$ into $P(x, Q)$ and require that ρ is such that the equation is satisfied at lowest order in Q, this process can be iterated to find r_i to the desired order in Q. The result of these calculations to fourth order in Q are summarized in Appendix A.

Using the series expansions of the roots we can Taylor expand the action as a power series in Q. This is slightly delicate because the complete elliptic integrals

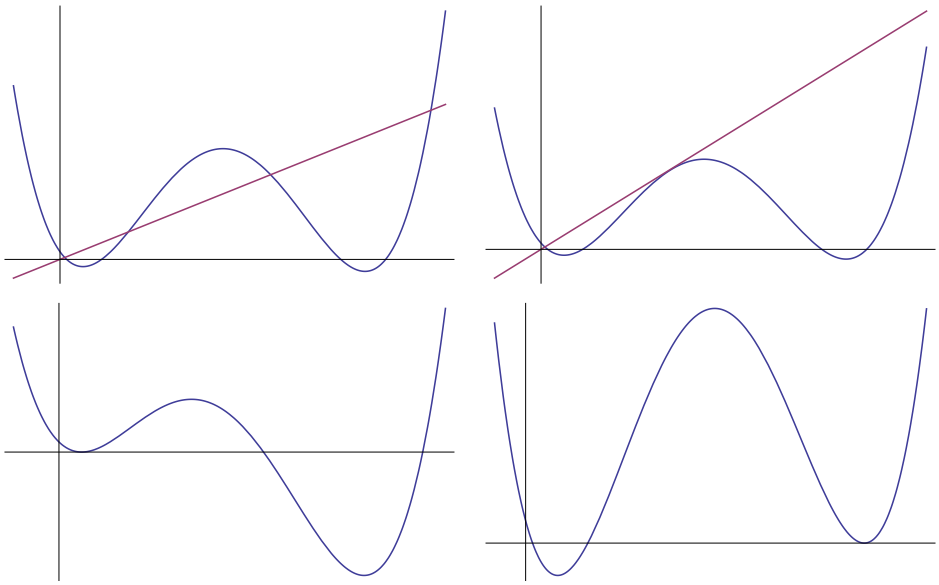


Figure 4.5: This figure summarizes our characterization of the shape space using the quartics $P_0(x)$ and $P(x)$, defined by equations (4.45) and (4.52). The straight lines are given by $y = 324Q^2x$ and the quartic curves by $y = P(x, 0) \equiv P_0(x)$. The upper left panel shows the generic case in which the fixed volume determines four distinct roots. The upper right panel shows the coalescence of the middle roots at maximum volume and $m = 0$. The lower two panels show the coalescence of other pairs of roots at zero volume and $m = 1$ (these are two distinct cases).

diverge logarithmically at $m = 1$, however, taking care to expand the logarithms to the proper order we find simple results. First we consider the case in which the A_r are equal with $A_r \equiv A_0$, so that $\bar{r}_1 = \bar{r}_2 \equiv 0$ and $\bar{r}_3 = \bar{r}_4 \equiv \tilde{r} = (2A_0)^2$. The action simplifies and the expansion yields,

$$I = 18gQ\left(\left[-1 - 2\frac{r_4}{(r_4 - \tilde{r})}\right] K(m) - 2\frac{\tilde{r}(r_4 - r_3)}{(r_4 - \tilde{r})(r_3 - \tilde{r})}\Pi(\alpha_4^2, m)\right) \tag{4.82}$$

$$\approx \sqrt{\tilde{r}}\pi + \frac{18Q}{\tilde{r}} \ln\left(\frac{9Q}{e\tilde{r}^{3/2}}\right)^3 + O(Q^2(\ln Q)) \tag{4.83}$$

$$= I(0) + (6e\tilde{r}^{1/2})\left(\frac{9Q}{e\tilde{r}^{3/2}}\right) \ln\left(\frac{9Q}{e\tilde{r}^{3/2}}\right), \tag{4.84}$$

where in the last equality we've recognized $\sqrt{\tilde{r}}\pi = 2\pi A$ as half of the symplectic area of the sphere and hence the action at zero volume squared, $I(0)$.

At lowest order then, the relation between I and Q can be inverted using Lambert's W function. Because this inverse function is uncommon, we briefly review the properties used in this work. The Lambert W is defined as the function which inverts

the relationship,

$$x = We^W, \tag{4.85}$$

yielding $W(x)$. The function $W(x)$ is plotted in Figure 3. Note that the function is

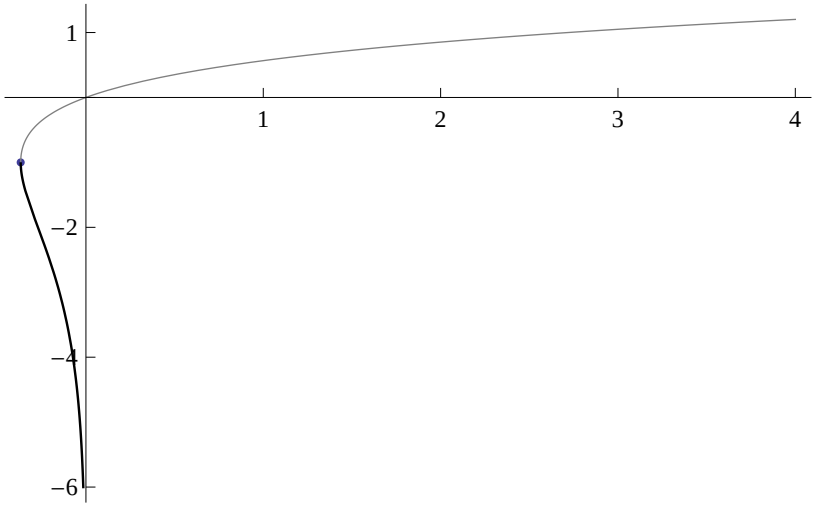


Figure 4.6: The real values of the Lambert W function

multivalued in the interval $[-1/e, 0]$, the upper branch (lighter shade in Figure 3) is conventionally taken to be the principle branch $W_0(x)$. However, for the expansion that we are interested in, we need the lower branch $W_{-1}(x)$ (darker shade in Figure 3). As x approaches zero from below $x \to 0_-$ the lower branch $W_{-1}(x)$ can be developed in the following series,

$$W_{-1}(x) \approx -\ln(-\frac{1}{x}) - \ln(\ln(-\frac{1}{x})) - \frac{\ln(\ln(-\frac{1}{x}))}{\ln(-\frac{1}{x})} + \cdots, \tag{4.86}$$

this is the series that we will need to complete our derivation of a lower bound on the volume spectrum.

Using these observations about the Lambert W function we have from (4.84),

$$Q \approx \frac{\tilde{r}(S - S_0)}{54 W_{-1}((S - S_0)/(6e\tilde{r}^{1/2}))}. \tag{4.87}$$

And expanding this solution for small $(S - S_0)/(6e\tilde{r}^{1/2})$ yields,

$$V \approx \frac{2\sqrt{\pi}}{3\sqrt{3}} A \left(\frac{1}{\sqrt{\ln\left(\frac{6eA}{\pi}\right) + \ln\left(\ln\left(\frac{6eA}{\pi}\right)\right) + \cdots}} \right), \tag{4.88}$$

in the case that the phase space is odd dimensional, and a very similar result in the even dimensional case. Both cases are plotted in Figure 4.7 the odd dimensional case being the middle curve and the even dimensional case the lowest curve. Note that the smallest Bohr-Sommerfeld eigenvalues in this plot are in much poorer agreement than for the case of Figure 4.7. This is due to the fact, discussed in section 4.4, that the space of shapes is no longer a differentiable manifold when we consider equal spins j_r, i.e. equal areas A_r. Roughly speaking it is like the sphere develops a cusp at its north or south poles or perhaps both depending on the values of the A_r.

4.7 Wavefunctions

In the last two section we have seen the utility of the Bohr-Sommerfeld approximation for finding the eigenvalues of the volume operator. Semiclassical techniques can also be used to find the volume wavefunctions. Before commencing the calculation we provide a quick recap of the geometrical WKB theory outlined in Chapter 2. Recall that the wavefunction of an observable $\hat{A}$ evaluated in the basis of an observable $\hat{B}$ is given by[7]

$$\langle b|a\rangle = \frac{(2\pi i)}{\sqrt{V_A V_B}} \sum_{\{p\}} |D_p|^{-1/2} \exp\{i[S_{Ap} - S_{Bp} - \frac{\mu_p \pi}{2}]\}. \tag{4.89}$$

The sum is over the set of intersections $\{p\}$ of the level sets $A = a$ and $B = b$ and V_A and V_B are the volumes of these level sets. The "amplitude determinant" D_p is given by $D_p = \{A, B\}|_p$ and μ_p is the Maslov. The action function S_{Ap} is given by the integral of the symplectic 1-form from a conventional point p_{A0} to the intersection point p along the A level set and similarly for S_{Bp}.

Applying this general theory to the calculation of the volume wavefunctions is straightforward. Take the A-states to be eigenstates of $\hat{Q}$, $|q\rangle$ with the corresponding classical function Q. These states can be expressed as wavefunctions in the basis $|k\rangle$ of $\mathcal{K}_4$ developed in section 4.3.1. The classical function, corresponding to the operator for which these are the basis states, is $\boldsymbol{A}$. In Figure 4.8 the intersection of a constant Q level set with a constant A levels set is depicted. The $|\boldsymbol{A}|$ level sets are lines of latitude on the shape space sphere. The intersections of these two contours fall into three classes: zero intersections which corresponds to a classically forbidden region; one intersection, at either the minimum or the maximum of the Q level set which corresponds to a caustic; and two intersections $\{p_1, p_2\}$ which corresponds to a classically allowed region. The classically allowed region can also be visualized in terms of the quartic $P_0(x)$ and is given by the x interval between the two intermediate intersections of the upper left hand panel in Figure 4.5.

[7]The symbol A is used here in the general sense of Chapter 2 and does not refer to the magnitude of the area vector $\boldsymbol{A}$.

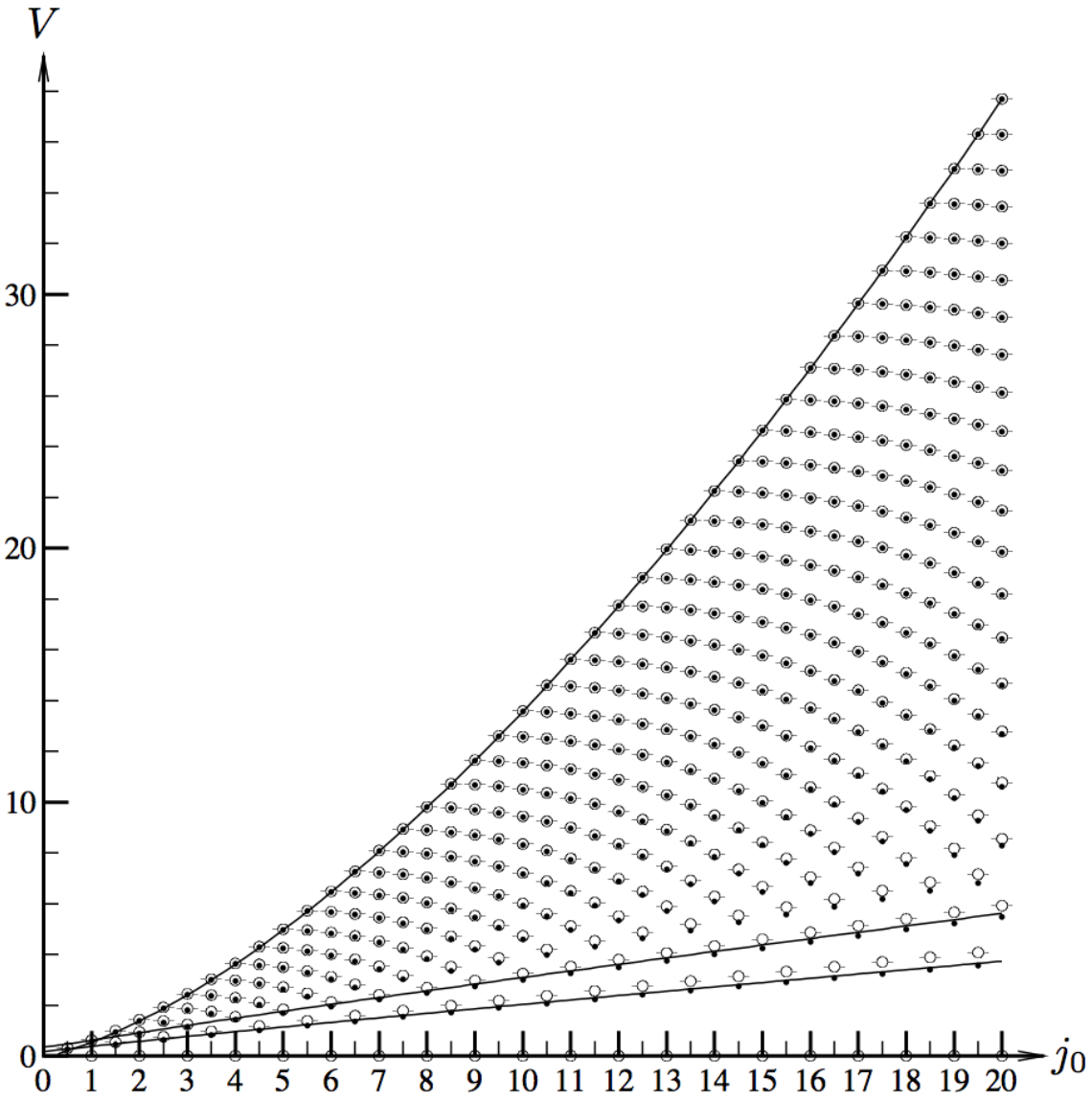


Figure 4.7: This figure compares the semiclassical scaling of the largest and smallest volume eigenvalues (dark lines) to the Bohr-Sommerfeld (dots) and loop gravity (circles) spectra. The four spins are equal in this case $j_1 = j_2 = j_3 = j_4 = j$ and this explains the poorer agreement of the spectra at small eigenvalues.

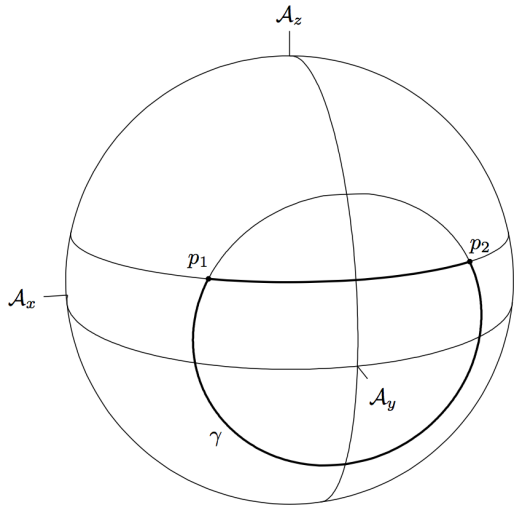


Figure 4.8: The intersections p_1 and p_2 of a constant Q and a constant A level set. The darkened contour γ is used to calculate the action loop integral $S_q - S_k$.

As often happens in problems with a discrete symmetry (here given by $\phi \to \pi - \phi$) the general formula (4.89) can be simplified. The manipulations that follow will separate the wavefunction into two pieces, one that depends on the conventional starting points p_{q_0} and p_{k_0} and one that is independent of all conventions. Here and below we use the eigenvalues $\{q, k\}$ to label level sets rather than the classical functions, the resulting notation is clearer. We will find below that $D_{p_1} = D_{p_2} \equiv D_0$ and so the amplitude can be pulled out of the sum. If an additional phase factor $\prod_p \exp\left\{i\left[\frac{S_{qp}-S_{kp}}{2} - \frac{\mu_p \pi}{4}\right]\right\}$ is pulled out and all the phases are collected into a single unknown factor η then (4.89) becomes,

$$\langle k|q\rangle = \frac{(2\pi)^{1/2}}{\sqrt{V_k V_q}} |D_0|^{-1/2} \eta \cdot 2\cos\left(\frac{S_q - S_k}{2} - \frac{\mu\pi}{4}\right), \tag{4.90}$$

where $S_q \equiv S_{q1} - S_{q2}$ is the integral from intersection 1 to intersection 2 and hence independent of p_{q_0}, and similarly for S_k. The difference $S_q - S_k$ is the loop integral around the darkened contour in Figure 4.8. The total maslov index $\mu \equiv \mu_1 - \mu_2$ is also independent of any conventions and so we have achieved the aforementioned separation; only η depends on the conventions set by p_{q_0} and p_{k_0}. In section 4.5 we assembled all of the necessary tools for explicitly calculating this formula, we are only left with a few short computations.

We begin with the amplitude factors: The amplitude determinant is given by (see (4.48) and (4.40)),

$$D_0 = \{A, Q\} = \frac{8\Delta\bar{\Delta}}{9A}\cos(\phi) = \frac{1}{18A}\sqrt{(16\Delta\bar{\Delta})^2 - (18AQ)^2}. \tag{4.91}$$

The last equality demonstrates that the amplitude is independent of ϕ. The level sets of A only depend on ϕ and hence as claimed above $D_1 = D_2 = D_0$. Because the level sets are both one dimensional the volumes V_k and V_q are just the periods of the level set. For the A evolution this is just $V_k = 2\pi$, while for the Q evolution this was calculated just below (4.56), $V_q = 18QK(m)$, with the elliptic parameter given by (4.55).

Next we calculate the action loop integral $S_q - S_k$. The contour of integration γ is the darkened curve in Figure 4.8. Both the A and the Q level sets are symmetric about the great circle arc $\phi = \pi/2$, the A level set being a small circle and the Q level set being given by $Q = 8\Delta\bar{\Delta}\sin\phi/(9A)$. This allows us to split γ in half and double the contribution of the contour with $\phi \geq \pi/2$,

$$S_q - S_k = \oint_\gamma A d\phi = 2\int_{9gK}^{\lambda(p_2)} A\frac{d\phi}{d\lambda}d\lambda + 2\int_{\phi(p_2)}^{\pi/2} A d\phi. \tag{4.92}$$

We treat these contributions one at a time. Let $\phi(p_2) \equiv \phi_0$ and observe that along the A level set, A is a constant so that

$$S_1 \equiv \int_{\phi_0}^{\pi/2} A d\phi = 2A(\frac{\pi}{2} - \phi_0). \tag{4.93}$$

The intersection point with $\phi \geq \pi/2$ is given by $\phi_0 = \arcsin 9AQ/(8\Delta\bar{\Delta})$ and here the branch of arcsin with $\phi \in [\pi/2, 3\pi/2]$ is understood.

The second contribution, along the Q level set, is a generalization of the action I calculated above and can once more be expressed in terms of elliptic integrals of the first and third kinds. Let $\lambda(p_2) \equiv \lambda_0$ then,

$$\begin{aligned} S_2 &\equiv 2\int_{9gK}^{\lambda_0} A\frac{d\phi}{d\lambda}d\lambda \\ &= 18gQ\left(\left(1 - \sum_i \frac{r_4}{r_4 - \bar{r}_i}\right)\frac{\lambda_0}{9g}\right. \\ &\qquad \left. - \sum_i \frac{\bar{r}_i(r_4 - r_3)}{(r_3 - \bar{r}_i)(r_4 - \bar{r}_i)}\Pi\left(\alpha_i^2, \mathrm{am}\left(\frac{\lambda_0}{9g}, m\right), m\right)\right) - I, \end{aligned} \tag{4.94}$$

where I is the Bohr-Sommerfeld action found in section 4.5 and given by (4.68). Due to the quantization condition $I = (n + 1/2)2\pi$ where n is the quantum level of the Q level set. The elliptic integrals Π are incomplete and depend on three parameters, α_i and m, as before, and also on the Jacobi amplitude of the intersection point $\mathrm{am}\,(\lambda_0/(9g), m)$. This amplitude is most easily evaluated by noting that if $\psi \equiv \mathrm{am}\,(\lambda_0/(9g), m)$ then,

$$\sin\psi = \sin\left(\mathrm{am}\left(\frac{\lambda_0}{9g}, m\right)\right) = \mathrm{sn}\left(\frac{\lambda_0}{9g}, m\right) = \sqrt{\frac{(x_0 - r_3)(r_4 - r_2)}{(x_0 - r_4)(r_3 - r_2)}} \tag{4.95}$$

where the last equality follows from (4.56). As usual $x \equiv A^2$ and so x_0 is evaluated on the constant A level set. Finally we need to calculate the overall phase and the maslov index.

To get a handle on the phase we return to the structure of the matrix elements $\langle k'|Q|k\rangle$ (see (4.31)).[8] Note that these matrix elements imply a recursion relation between wavefunctions in the $|k\rangle$ basis:

$$q\langle k|q\rangle = -ia_i\langle k-1|q\rangle + ia_{i+1}\langle k+1|q\rangle \tag{4.96}$$

where the a_i are real coefficients whose exact expression is unnecessary for what follows. Looking to the lowest weight state $|k_{\min}\rangle$ this recursion becomes, $\langle k_{\min}+1|q\rangle = -i\frac{q}{a_{i+1}}\langle k_{\min}|q\rangle$. Iterating this procedure shows that the overall phase depends on k as $\exp ik\pi/2$. The remaining phase is purely conventional (depending on the choice of p_{q_0} and p_{k_0}) and we adopt the convention that the lowest weight element $\langle k_{\min}|q\rangle$ is real and positive. Finally, calculating the Maslov index one finds that it is $\mu = 1$.

Putting all of these elements together we obtain the asymptotic wavefunction of the volume operator,

$$\psi_q(k) = \frac{1}{\sqrt{gK(m)}}\frac{2\sqrt{A}}{|((16\Delta\bar{\Delta})^2 - (18AQ)^2)^{1/4}|}\eta\cos\left(\frac{1}{2}\left(S_1+S_2\right) - \frac{\pi}{4}\right), \tag{4.97}$$

with S_1 and S_2 given by equations (4.93) and (4.94).

In Figure4.9 we plot an exact eigenvector of $\langle k'|Q|k\rangle$ calculated numerically along with the WKB approximation of the wavefunction. For the case pictured here the eigenvector and wavefuntion are real but this not the case in general.

4.8 The 5-valent volume evolution: a beginning

All of the volume operator proposals discussed in section 4.3 agree at the level of the 4-valent node, more precisely they are proportional. However, for 5- and higher valent nodes these operators begin to disagree. In this section we initiate an investigation of the semiclassical volume evolution of the 5-valent node. By volume evolution we mean the dynamical evolution of the phase space variables when the volume V_{Pol} of the 5-valent node is taken as Hamiltonian. In particular we are interested in discovering whether this evolution is chaotic (the generic case in 2 dimensional configuration spaces) or integrable and if it is integrable in finding what function commutes with the classical volume.

These two possibilities have important consequences for loop gravity. If the classical volume generates a chaotic flow then the corresponding quantum spectrum will generically be non-degenerate and the volume eigenvalue continues to act as a good

[8] This argument is inspired by [115].

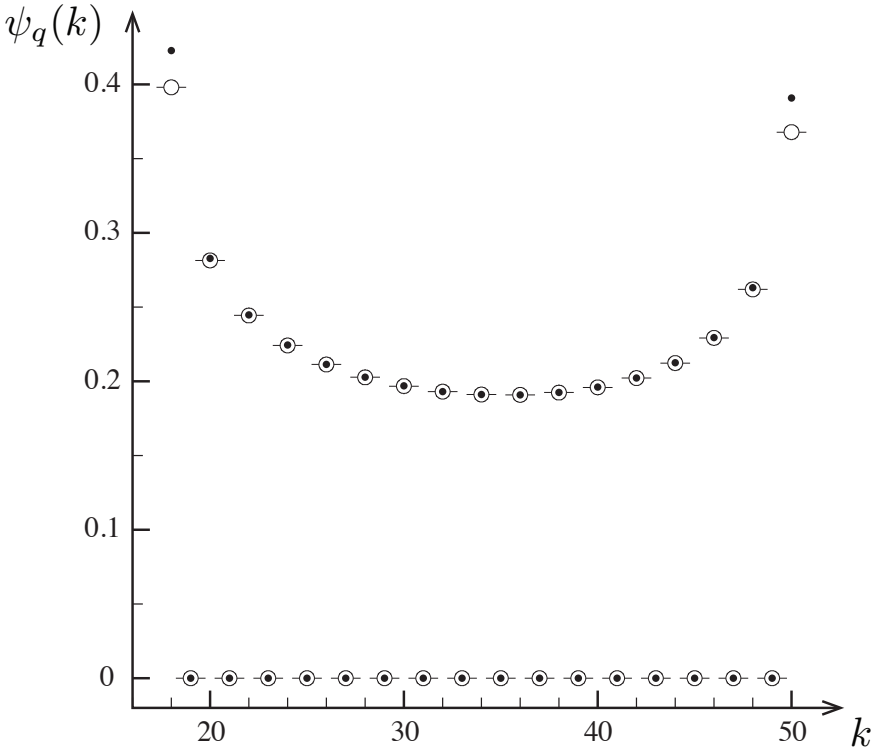


Figure 4.9: Comparison of the exact and WKB volume wavefunctions for the spins $j_1 = j_2 = j_3 = 25$ and $j_4 = 43$. The wavefunctions pictured are for the level set $q = 0$. This is an eigenvalue of $\hat{Q}$ because the intertwiner space is odd dimensional, $\dim \mathcal{K}_4 = 33$.

label for spin network states. On the other hand if the volume flow is classically integrable then the degeneracy of the corresponding quantum spectrum will have to be lifted by another observable. Both of these outcomes will impact the direction of future research into the volume operator.

This section reports on recently initiated research along these lines. We have not yet discovered whether the volume evolution is integrable or chaotic however we have removed most of the hurdles to answering this question. One of these hurdles is that there is no general analytic solution to the Minkowski reconstruction problem for collections of area vectors consisting of more than four vectors. This problem has proven to be numerically tractable [109, 128] but these numerical studies are impractical for studying the volume evolution. In this section we find a complete analytical solution for the Minkowski reconstruction in the case of 5 area vectors. It turns out that the volume of the polyhedron plays a role in this construction and so as a side benefit we obtain an explicit formula for the volume in terms of the area vectors as well.

4.8.1 Volume and adjacency of the triangular prism

In order to perform analyses like those in the tetrahedral case we would like to find the volume of the triangular prism in terms of its area vectors. In this section we use "triangular prism" in a generalized sense intended to mean the generic convex polyhedron with 5 faces. Two of the faces of this polyhedron are always triangular and they are opposite in the sense that they have no edges in common. Given five area vectors that satisfy,

$$\boldsymbol{A}_1 + \boldsymbol{A}_2 + \boldsymbol{A}_3 + \boldsymbol{A}_4 + \boldsymbol{A}_5 = 0, \tag{4.98}$$

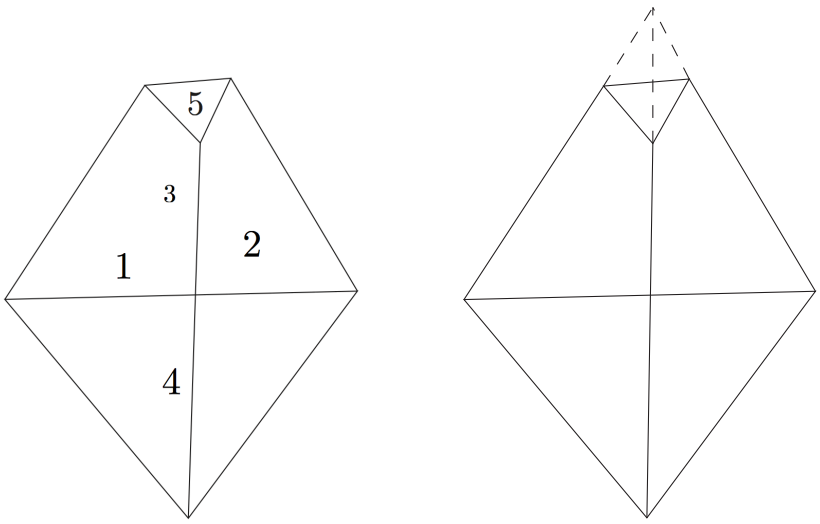
the Minkowski theorem guarantees that there is a unique convex polyhedron corresponding to these vectors. Our strategy for finding the volume of this polyhedron in terms of its area vectors will be to try and connect this polyhedron with the tetrahedral case studied above.

One of these polyhedra and the convention for its face labeling are shown in the left panel of Figure 4.10. The right panel of Figure 4.10 shows a completion of this triangular prism into a large tetrahedron made up of the original triangular prism plus a small tetrahedron. Let α, β and γ be real numbers greater than 1 such that the face area vectors of the large tetrahedron are $\alpha\boldsymbol{A}_1, \beta\boldsymbol{A}_2, \gamma\boldsymbol{A}_3$ and $\boldsymbol{A}_4$, then this tetrahedron satisfies the closure condition,

$$\alpha\boldsymbol{A}_1 + \beta\boldsymbol{A}_2 + \gamma\boldsymbol{A}_3 + \boldsymbol{A}_4 = 0. \tag{4.99}$$

Similarly the area vectors of the small tetrahedron are given by $(\alpha-1)\boldsymbol{A}_1$, $(\beta-1)\boldsymbol{A}_2$ and $(\gamma-1)\boldsymbol{A}_3$ and satisfy,

$$(\alpha-1)\boldsymbol{A}_1 + (\beta-1)\boldsymbol{A}_2 + (\gamma-1)\boldsymbol{A}_3 - \boldsymbol{A}_5 = 0. \tag{4.100}$$



(a) Triangular prism with area labels. The top triangle is labelled 5 and the base trianlge 4.

(b) Extension of prism into a tetrahedron

Figure 4.10: The triangular prism.

Dotting $(\boldsymbol{A}_2 \times \boldsymbol{A}_3)$ into (4.99) yields,

$$\alpha = -\frac{\boldsymbol{A}_2 \cdot (\boldsymbol{A}_3 \times \boldsymbol{A}_4)}{\boldsymbol{A}_1 \cdot (\boldsymbol{A}_2 \times \boldsymbol{A}_3)}. \tag{4.101}$$

Equations for β and γ are derived in a similar manner,

$$\beta = \frac{\boldsymbol{A}_1 \cdot (\boldsymbol{A}_3 \times \boldsymbol{A}_4)}{\boldsymbol{A}_1 \cdot (\boldsymbol{A}_2 \times \boldsymbol{A}_3)}, \qquad \gamma = -\frac{\boldsymbol{A}_1 \cdot (\boldsymbol{A}_2 \times \boldsymbol{A}_4)}{\boldsymbol{A}_1 \cdot (\boldsymbol{A}_2 \times \boldsymbol{A}_3)}. \tag{4.102}$$

Note that although α, β and γ must be greater than 1 for the closure relation (4.99) to be satisfied, they are well defined parameters for all area vectors. Once again the volume of a tetrahedron can be derived from its area vectors via,

$$V = \frac{\sqrt{2}}{3}\sqrt{|A_r \cdot (A_s \times A_t)|} \equiv \frac{\sqrt{2}}{3}\sqrt{|W_{rst}|}, \tag{4.103}$$

where $r \neq s \neq t$ and we've introduced the shorthand $W_{rst} \equiv \boldsymbol{A}_r \cdot (\boldsymbol{A}_s \times \boldsymbol{A}_t)$. This allows us to express the volume of the prism V_p as the difference of volumes of the two tetrahedra,

$$V_p = \frac{\sqrt{2}}{3}\sqrt{|W_{123}|}\left(\sqrt{\alpha\beta\gamma} - \sqrt{(\alpha-1)(\beta-1)(\gamma-1)}\right). \tag{4.104}$$

This formula for the volume is only valid when the adjacency relations of the faces are as depicted in Figure 4.10(a) and when the closure is that of the tetrahedron in Figure 4.10(b). When the adjacency and closure are otherwise, another similar but different formula is valid. For this reason it will be essential to determine the adjacency of the faces and the minimal closure orientation from the given area vectors in order to calculate the volume and its Hamiltonian evolution.

The Minkowski theorem proves existence and uniqueness but does not specify how to construct the convex polyhedron it guarantees. This reconstruction is also greatly aided by determination of the adjacency relations of the polyhedron's sides, as will be shown below. The adjacency is determined from the given vectors as follows: if instead of the closure relation (4.99) we assume instead a different closure, for example the one with,

$$\alpha_2 \boldsymbol{A}_1 + \beta_2 \boldsymbol{A}_2 + \gamma_2 \boldsymbol{A}_3 + \boldsymbol{A}_5 = 0, \tag{4.105}$$

then,

$$\alpha_2 = -\frac{\boldsymbol{A}_2 \cdot (\boldsymbol{A}_3 \times \boldsymbol{A}_5)}{\boldsymbol{A}_1 \cdot (\boldsymbol{A}_2 \times \boldsymbol{A}_3)} = 1 + \frac{\boldsymbol{A}_2 \cdot (\boldsymbol{A}_3 \times \boldsymbol{A}_4)}{\boldsymbol{A}_1 \cdot (\boldsymbol{A}_2 \times \boldsymbol{A}_3)} = 1 - \alpha, \tag{4.106}$$

where in the second equality the assumed closure $\boldsymbol{A}_1 + \cdots \boldsymbol{A}_5 = 0$ was used and in the third the definition of α. Similarly, β_2 and γ_2 can be expressed in terms of β and γ yielding, $\beta_2 = 1 - \beta$ and $\gamma_2 = 1 - \gamma$. This demonstrates that these two closure relations are mutually exclusive, if $\alpha, \beta, \gamma > 1$ then $\alpha_2, \beta_2, \gamma_2 < 0$.

The adjacency of a triangular prism is determined by which two of its faces are triangular (each of the quadrilateral faces is adjacent to all other faces). Thus there are $\binom{5}{2} = 10$ different adjacencies. If one further specifies which of the triangular sides the prism closes on, as in Figure 4.10(b), then there are a total of 20 configurations consisting of adjacency and tetrahedral closure. On the other hand to specify the tetrahedral closure condition algebraically you need to choose 4 out of the 5 area vectors and then to choose 3 of these vectors to scale to close the tetrahedron, this is $\binom{5}{4} \cdot \binom{4}{3} = 20$ different closure relations. This shows that all twenty cases of adjacency plus tetrahedral closure information can be covered by writing down the appropriate tetrahedral closure relation with its α_i, β_i and $\gamma_i, (i = 1, \cdots, 20)$. The exhaustive list of these parameters is displayed in Appendix B along with the algebraic relations to the original α, β and γ parameters introduced above. A remarkable feature of this list is that if one assumes that $\alpha, \beta, \gamma > 1$ it follows that 13 out of the twenty cases are excluded.

This exclusion of cases has an appealing geometrical interpretation. The assumption that $\alpha, \beta, \gamma > 1$ implies that it may be possible to construct a polyhedron with the adjacencies of Figure 4.10(a). The excluded cases are precisely those that cannot be reached by a Schlegel move (defined momentarily) from this adjacency. A Schlegel diagram is a planar graph obtained by choosing one of the faces of a convex polyhedron and projecting the rest of the polyhedron into this face. We define a Schlegel move as a merging and re-splitting of two of the vertices of this graph that preserves

the number of edges of the graph. The Schlegel move taking the prism with faces 5 and 4 opposite to that with 1 and 2 opposite can be visualized using Figure 4.10. Imagine the edge connecting the two vertices in the foreground shrinking until the two vertices merge, at this point the figure has become a pyramid over a a quadrilateral base. Now, let this vertex split apart again but this time with the opposite orientation, causing faces 1 and 2 to become opposite triangles. This is the Schlegel move connecting the 54 prism to the 12 prism. There is no Schlegel move taking the 54 prism to a 52 prism that does not degenerate an entire face of the polyhedron. Algebraically this corresponds to the fact that the $\alpha\beta\gamma$-parameters of a 52 prism are mutually exclusive with those of a 54 prism. This is the sense in which these cases are mutually excluded.

Conversely the cases that are consistent with $\alpha, \beta, \gamma > 1$ are those which can be reached by a Schlegel move from a 54-prism. The number of adjacencies that are potentially constructible can be further reduced by ordering α, β and γ. For example, if we assume $\alpha > \beta > \gamma$ in addition to $\alpha, \beta, \gamma > 1$ then there are only two cases that are consistent, the 54- and 12-prisms (cases 1. and 20. of appendix B). Each of the six orderings of α, β and γ corresponds to consistency of one of the 6 Schlegel diagrams reachable from a 54-prism. Thus the problem of which adjacency and tetrahedral closure is constructible has been reduced to distinguishing between a single pair of cases that can be connected by a Schlegel move.

The distinction between these final two cases can also be made using α, β and γ. Consider the case discussed above in which $\alpha > \beta > \gamma > 1$. The transition from the 54-prism to the 12-prism occurs when the two vertices in the foreground of Figure 4.10 have merged into a single vertex. This results in a pyramidal configuration. Just as before the faces 1, 2 and 3 can be extended to close this pyramid into a tetrahedron. However, now it is also possible to extend faces 3, 4 and 5 to close it into a second distinct tetrahedron. Using these two extensions the volume of the pyramidal polyhedron can be calculated in two distinct ways. Setting these two volume formulae equal yields an algebraic relation between α, β and γ:

$$\gamma = \frac{\alpha\beta}{\alpha + \beta - 1}. \tag{4.107}$$

When $\gamma > (\alpha\beta)/(\alpha+\beta-1)$, the 54-prism is constructible and when $\gamma < (\alpha\beta)/(\alpha+\beta-1)$, the 12-prism is constructible. The proof is elementary but would be a distraction to include here. This discussion provides a complete solution to the problem of the face adjacency by calculating and ordering the twenty sets of ratios of triple products enumerated in Appendix B.

4.8.2 Explicit Minkowski Reconstruction

Once the adjacency and tetrahedral closure relation are known the reconstruction of the polyhedron from the given area vectors proceeds in a simple fashion. Let

h_r $(r = 1, \ldots, 5)$ be five heights that measure the perpendicular distance from the face determined by $\boldsymbol{A}_r$ to an arbitrary origin. Assume that the 54-prism is the constructible one and choose the origin to be at the vertex where the planes 1, 2 and 4 meet, then $h_1 = h_2 = h_4 = 0$. The vector pointing to the intersection of three planes can be expressed in terms of the normals to the planes $\hat{n}_r, \hat{n}_s$ and $\hat{n}_t$ by,

$$\boldsymbol{x}_{rst} = \frac{1}{|\hat{n}_r \cdot (\hat{n}_s \times \hat{n}_t)|} \left[h_r(\hat{n}_s \times \hat{n}_t) + h_t(\hat{n}_r \times \hat{n}_s) + h_s(\hat{n}_t \times \hat{n}_r) \right], \tag{4.108}$$

as can easily be confirmed by dotting in $\hat{n}_r, \hat{n}_s$ and $\hat{n}_t$. This formula can be used to find the edge lengths of face 4 and with the convention that $h_1 = h_2 = h_4 = 0$ the edge lengths are all proportional to h_3. Let these three edge lengths be $h_3 l_1, h_3 l_2$ and $h_3 l_3$ where the dimensionless "lengths" (l_1, l_2, l_3) are completely determined by the given $\boldsymbol{A}_r$ and the formula (4.108). Then using Heron's formula, $A_4 = h_3^2 \Delta(l_1, l_2, l_3)$ and this can be solved for h_3,

$$h_3 = \sqrt{\frac{A_4}{\Delta(l_1, l_2, l_3)}}. \tag{4.109}$$

Finally, h_5 can be extracted from the relation of the volume to the heights,

$$V_p = \frac{1}{3}(A_3 h_3 + A_5 h_5) \tag{4.110}$$

and the value of the volume determined by (4.104). This completes the Minkowski reconstruction.

4.8.3 Outlook for the 5-valent volume

Of course, the Minkowski reconstruction is only tangential to the goal of understanding the volume evolution of the triangular prism. The real accomplishment of this section is a complete characterization of the adjacency of the faces of the polyhedron in terms of the given area vectors $\boldsymbol{A}_r$. Once the adjacency is known there is a definite formula for the volume of the polyhedron. All of these formulae are similar to the 54 prism formula displayed above, (4.104). This formula is a complicated nonlinear function of the shape space coordinates μ_k and ϕ_k of section 4.2 and there is little hope of an analytical solution for the volume evolution. However, we have been able to lift the problem into the relevant Schwinger space and perform a numerical integration using a symplectic leap frog integrator. This integrator does an excellent job of preserving the face areas A_r as is required of the volume evolution (they are Casimirs of the Poisson bracket). Using well developed techniques in the numerical study of chaos we should be able to answer the question of whether the classical volume evolution of the triangular prism is chaotic or integrable in the near future.

Before proceeding to the conclusion of this chapter we return to the issue of cylindrical consistency, which was discussed in the literature review at the end of the first part of section 4.3. The classical volume of a triangular prism does reduce to the volume of a tetrahedron as you take one of the face areas of the prism to zero but it does this in a non-trivial fashion. The limiting of either triangular face area to zero is straightforward, the three vertices of the triangle merge and leave a tetrahedron. If the limiting of a quadrilateral face worked in the same manner all four of its vertices would merge and the polyhedron would degenerate to a triangle. Instead as you decrease the area of a quadrilateral face, say face 2, eventually you change the adjacency of the faces and face 2 becomes triangular. After this point the limiting proceeds once more in the intuitive fashion. This shows that the polyhedral picture emphasized in this chapter accomplishes cylindrical consistency in a subtle manner.

4.9 Conclusions

At the Planck scale, a quantum behavior of the geometry of space is expected. Loop gravity provides a specific realization of this expectation: it predicts a granularity of space with each grain having a quantum behavior. In this chapter we have presented a new independent road to the granularity of space and the computation of the spectrum of the volume. The derivation is based on semiclassical arguments applied to the simplest model for a grain of space, a Euclidean tetrahedron, and is closely related to Regge's discretization of gravity and to more recent ideas about general relativity and quantum geometry [8, 9]. The spectrum has been computed by applying Bohr-Sommerfeld quantization to the volume of a tetrahedron seen as an observable on the phase space of shapes.

There is remarkable quantitative agreement of the spectrum calculated here and the spectrum of the volume in loop gravity. This result lends further credibility to the intricate derivation of the volume spectrum in loop gravity, showing that it matches with our elementary semiclassical approach. The semiclassical methods of this chapter provide a new understanding of many aspects of the rich structure of the volume spectrum in loop gravity and the explicit formulas open new avenues for analytical exploration. This is important because a deep understanding of the spectra of geometrical operators provides fertile ground for developing phenomenological tests of loop gravity [129]. The same methods can be applied to other geometrical operators, as well as to the alternative versions of the volume operator considered in the literature.

In loop gravity, the discrete spectra of geometrical observables provide a physical Planck-scale cut-off that renders the theory finite in the ultraviolet [1, 2, 130, 131]. An important question is whether there exists a volume gap, that is a discrete gap, above zero, in the volume spectrum for all spins. All of the results on small eigenvalues

presented here support the existence of a volume gap for the 4-valent case. An intriguing aspect of the investigation of small eigenvalues of the volume operator was the appearance of the unusual series (4.88) describing the scaling of the small eigenvalues. The semiclassical corrections, leading to a log modified linear scaling, would have been very difficult to guess.

We have also computed the eigenvectors of the volume in a WKB expansion. The earliest work on the volume operator explored in this chapter [111, 112] was performed with applications to atomic physics in mind. The idea was that the volume eigenstates $|q\rangle$ provide a democratic coupling of three angular momenta that have nice properties under the permutations of the three momenta. It would be interesting to see if our explicit semiclassical formulas for the volume wavefunctions could give rise to new applications in atomic physics.

When $F > 4$, the phase space $\mathcal{P}_F$ has dimension greater than two and generic Hamiltonians will generate chaotic dynamics [132]. A preliminary analysis of the case $F = 5$ indicates that, while the volume orbits may be chaotic, the dynamics can still be practically investigated numerically. This opens up the exciting possibility for exploring quantum chaos in the volume spectrum of loop gravity.

Chapter 5

The $9j$-symbol and asymptotics of complex spin networks

5.1 Introduction

The asymptotic behavior of spin networks has played a significant role in simplicial approaches to quantum gravity.[1] Indeed, the field began with the observation that the Ponzano-Regge action [11] for the semiclassical $6j$-symbol is related to the Einstein-Hilbert action of a tetrahedron in 3-dimensional gravity in the Regge formulation. To explain the meaning of our asymptotic result for the $9j$-symbol some analogies with the Ponzano-Regge formula for the $6j$-symbol are useful and so we describe their result now. The Ponzano-Regge formula is

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{Bmatrix} = \frac{1}{\sqrt{12\pi V}} \cos\left(\sum_{i=1}^{6} J_i\theta_i + \frac{\pi}{4}\right) \tag{5.1}$$

This equation holds in an asymptotic sense where all of the j's are taken to be large.

In the classically allowed region, the Ponzano-Regge formula associates a given $6j$-symbol with a real tetrahedron whose six edge lengths are $J_i = j_i + 1/2$, where the six j's are those appearing in the $6j$-symbol. More precisely, there are two tetrahedra, related by spatial inversion, that is, time-reversal. Except for flat configurations, the two tetrahedra are not related by proper rotations in $SO(3)$. We recall that time-reversal, not parity, inverts the direction of angular momentum vectors. The two tetrahedra correspond to the two stationary phase points of the $6j$-symbol, which make contributions to the asymptotic expression that are complex conjugates of each other. The result is the real cosine term in the Ponzano-Regge formula. One can say

[1] The main body of this chapter is taken directly from the paper [70], I thank my coauther R. G. Littlejohn and the publisher for permission to reproduce this work here. Because many small changes and a few new insights have been added I offer this footnote instead of using block quotation to mark quoted text.

that semiclassically the $6j$-symbol is a superposition of two amplitudes, corresponding to a tetrahedral geometry and its time-reversed image, that produce oscillations in the result.

We shall use lower case j's for quantum numbers, and capital J's for the lengths of the corresponding classical vectors. These are always related by $J_i = j_i + 1/2$. The $1/2$ is a Maslov index and the manner in which it arises in this context is explained in Aquilanti *et al* [37]. The tetrahedron associated to a $6j$-symbol is determined not just by its edge lengths, in general there are more than two tetrahedra with six given edge lengths, but also by the triangle inequality relations between various triples of j's. For the symbol above the triangle related triads are $\{j_1 j_2 j_3, j_2 j_4 j_6, j_1 j_5 j_6, j_3 j_4 j_5\}$. The edge lengths associated to these triads make up the four faces of the tetrahedron and this combined data uniquely determines it up to rotation and time reversal. The angles θ_i are the external dihedral angles, that is the angles between outward pointing face normals, about the edge with length J_i. These angles are equal for a tetrahedron and its time reversed image. They are uniquely determined by the edge lengths and face relations of the tetrahedron.

The relationship between the $6j$-symbol and tetrahedral geometry led Ponzano and Regge to introduce a state sum (discrete path integral) for 3-dimensional simplicial gravity. The $6j$-symbols are used as the weights of the tetrahedra in the generating function. This early insight has developed into the covariant spin foam approach discussed briefly in Chapter 1.

In this chapter we present the generalization of the Ponzano-Regge formula to the Wigner $9j$-symbol, as well as some material relevant for the asymptotics of arbitrary spin networks. The $9j$-symbol is the next most complicated spin network after the $6j$-symbol, with features that are found in all higher spin networks. In this chapter we present only the asymptotic formula itself for the $9j$-symbol and some salient facts surrounding it. We defer a derivation and deeper discussion of the formula.

Our derivation has quite a few steps, and some of them at this point are supported by numerical evidence only. Thus, we do not now have a rigorous derivation of our result. We believe it is correct, however, on the basis of direct numerical comparisons with the exact $9j$-symbol, the fact that our formula obeys all the symmetries of the exact $9j$-symbol, and the plausibility and numerical support for the conjectures involved in the parts of the derivation currently lacking proofs. The proofs do not seem difficult, and we hope to fill in the gaps in our future work.

Although most of the papers cited above have dealt with the asymptotics of specific spin networks, usually there are special values of the angular momenta that are used. For example, the $10j$-symbol involves balanced representations of $SO(4)$, which means that some pairs of j's are equal, while the $9j$-symbols that appear in LQG fusion coefficients have two columns in which one quantum number is the sum of the other two. In addition, j's are sometimes set equal because this is regarded as the most interesting regime from a physical standpoint.

As a result, the spin networks that have been studied tend to fall on caustics where

the asymptotic behavior is not generic. At such points, the value of the spin network (the wave function) is not oscillatory in a simple sense, instead it has the form of a diffraction catastrophe [133]. In addition, the wave function scales as a higher (less negative) power of the scaling parameter (effectively, $1/\hbar$). This type of behavior has been noted in several places in the quantum gravity literature, although as far as we can tell no one has noted that it is related to standard caustic and catastrophe types. In this chapter we give a rather complete picture of the $9j$-symbol for all possible parameters in the classically allowed region, including all phases and Maslov indices. We also indicate the subsets upon which the behavior is nongeneric and described by various types of caustics. We believe that this is the first time that such information has been available for any spin network more complicated than the $6j$-symbol.

Another reason for interest in the $9j$-symbol is that it is the nontrivial part of the Clebsch-Gordan coefficient for $SO(4)$.

Basic references on the Wigner $9j$-symbol include [99, 76, 134, 101]. Recent work on the $9j$-symbol has included new asymptotic forms when some quantum numbers are large and others small [135, 136]. We also note the use of $SU(2)$ spin networks in quantum computing [137].

In section 5.2 we present the asymptotic formula for the $9j$-symbol and draw comparisons with the Ponzano-Regge formula to introduce its geometrical content. A detailed explanation of the notation follows in later sections. In section 5.3 we present general rules for converting spin networks into surfaces composed of oriented edges and oriented triangles, and illustrate them for the $9j$-symbol. In section 5.4 we explain how the geometrical objects (pieces of oriented surfaces) corresponding to the $9j$-symbol can be constructed in 3-dimensional space. In section 5.5 we explain the configuration space of the $9j$-symbol and the classically allowed subset thereof. In section 5.6 we define the amplitude of the asymptotic formula and discuss the manifolds (the caustics) upon which it diverges as well as the diffraction catastrophes that replace the simple asymptotic form in the neighborhood of the caustics. In section 5.7 we explain the phase of the semiclassical approximation, a generalization of the Ponzano-Regge action that requires careful definitions of dihedral angles. In section 5.8 we show that the asymptotic formula correctly obeys the symmetries of the $9j$-symbol. Finally, in section 5.9 we present recent results on extending this analysis to higher $3nj$-symbols and conclusions.

5.2 The asymptotic formula

The asymptotic expression for the $9j$-symbol is

$$\left\{ \begin{array}{ccc} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \\ j_7 & j_8 & j_9 \end{array} \right\} = A_1 \cos S_1 + A_2 \sin S_2, \tag{5.2}$$

where $A_{1,2}$ are positive amplitudes, $S_{1,2}$ are phases, and each term is roughly similar to the single term in the Ponzano-Regge formula for the $6j$-symbol. More precisely

$$A = \frac{1}{4\pi\sqrt{V_{124}V_{542} - V_{451}V_{215}}}, \tag{5.3}$$

with $V_{ijk} \equiv \boldsymbol{J}_i \cdot (\boldsymbol{J}_j \times \boldsymbol{J}_k)$ and the subscripts 1,2 indicate that A is evaluated on two different geometries explained below. The phases are given by

$$S = \sum_{i=1}^{9} J_i \theta_i \tag{5.4}$$

with the angles θ_i defined below. The right hand side of (5.2) is the leading term in an asymptotic expansion in powers of $1/k$ of the $9j$-symbol when all nine j's are scaled by a positive factor k that is allowed to go to infinity (k plays the role of $1/\hbar$ in the asymptotic expansion). The k's are suppressed in (5.2), but the expression on the right scales as $1/k^3$. Equation (5.2) applies only in the classically allowed region. We do not present the analog of (5.2) in the classically forbidden region.

Equation (5.2) breaks down near caustics, where the $9j$-symbol scales with a higher (less negative) power of k than $1/k^3$. In the neighborhood of caustics, the $9j$-symbol is approximated by diffraction catastrophes, including the fold and hyperbolic and elliptic umbilic. These are discussed more fully in section 5.6.

As with the $6j$-symbol the asymptotics of the $9j$-symbol are determined by geometries that are fixed by the values of the nine j's. In the case of the $9j$-symbol in the classically allowed region, there are four geometrical figures associated with a given set of nine j's, consisting of two pairs related by time-reversal. The four geometrical figures correspond to the four real stationary phase points of the $9j$-symbol. Each pair of figures is associated with an "admissible" root (defined momentarily) of a certain quartic equation. There are two admissible roots in the classically allowed region, labeled 1 and 2, corresponding to the two terms in (5.2). Each trigonometric term in (5.2) consists of an exponential and its complex conjugate, corrresponding to a geometrical figure and its time-reversed image. One can say that semiclassically the $9j$-symbol is a superposition of four amplitudes corresponding to four geometries, consisting of two pairs of a geometry and its time-reversed image. We now explain these geometries and how they are specified by the nine j's that appear in the symbol.

5.3 Triangles, orientations and geometries

The $9j$-symbol specifies the lengths $J_i = j_i + 1/2$ of nine classical angular momentum vectors $\boldsymbol{J}_i$ but not their directions. Therefore we inquire as to how the directions may be determined, and geometrical figures constructed out of the resulting vectors.

Actually, it is convenient to double this set and speak of 18 classical vectors $\boldsymbol{J}_i$, $\boldsymbol{J}'_i$, $i = 1, \ldots, 9$. A doubling of this kind was introduced by Roberts [17], who gave

a highly symmetrical way of writing the $6j$-symbol as a scalar product in a certain Hilbert space. Although Roberts only worked with the $6j$-symbol, his method is easily generalized to an arbitary spin network. Ponzano and Regge [11] also gave hints that doubling of angular momentum vectors are important in the asymptotic analysis of spin networks.

We now describe rules that take an arbitary spin network (with at most trivalent vertices) and transcribe it into relations among a doubled set of classical angular momentum vectors, defining a set of oriented triangles and oriented edges of a geometrical figure. We exemplify these rules only in the case of the $9j$-symbol, but they are easily applied to any spin network. The reader may find it illuminating to apply our rules to the $6j$-symbol, starting with the usual spin network (the Mercedes graph). Figure 5.1 illustrates the spin network of the $9j$-symbol. See also Fig. 18.1 of Yutsis *et al* [106].

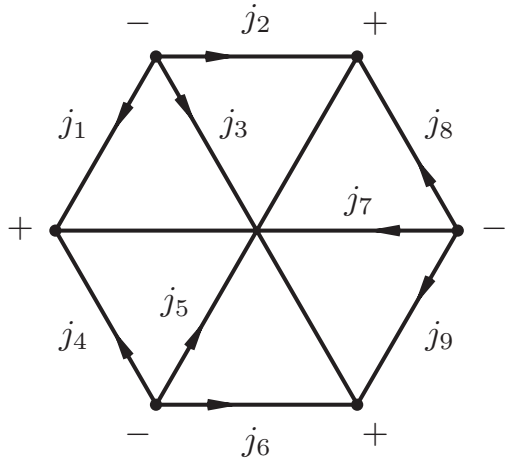


Figure 5.1: The spin network for the $9j$-symbol.

Each edge of the spin network, labeled by j_i, is associated with two classical angular momentum vectors $\boldsymbol{J}_i$ and $\boldsymbol{J}'_i$ that are required to satisfy

$$|\boldsymbol{J}_i| = |\boldsymbol{J}'_i| = J_i = j_i + 1/2 \tag{5.5}$$

and

$$\boldsymbol{J}_i + \boldsymbol{J}'_i = 0. \tag{5.6}$$

Vectors $\boldsymbol{J}_i$ and $\boldsymbol{J}'_i$ have the same length and point in opposite directions.

Each node of the spin network, where three edges meet, corresponds to three vectors that add to zero. The three vectors are associated with the three edges. If the arrow on an edge ending at the node is pointing away from the node, then the

angular momentum vector is unprimed; if it is pointing toward the node, then the vector is primed. This rule applied to Fig. 5.1 gives

$$\begin{aligned} \boldsymbol{J}_1 + \boldsymbol{J}_2 + \boldsymbol{J}_3 &= 0, \qquad & \boldsymbol{J}'_1 + \boldsymbol{J}'_4 + \boldsymbol{J}'_7 &= 0, \\ \boldsymbol{J}_4 + \boldsymbol{J}_5 + \boldsymbol{J}_6 &= 0, \qquad & \boldsymbol{J}'_2 + \boldsymbol{J}'_5 + \boldsymbol{J}'_8 &= 0, \\ \boldsymbol{J}_7 + \boldsymbol{J}_8 + \boldsymbol{J}_9 &= 0, \qquad & \boldsymbol{J}'_3 + \boldsymbol{J}'_6 + \boldsymbol{J}'_9 &= 0. \end{aligned} \tag{5.7}$$

These are a set of classical triangle relations, one for each node of the spin network. In the case of the $9j$-symbol, they are obviously related to the rows and columns of the symbol.

Although the vector addition in (5.7) is commutative, we agree to write the vectors in each equation in counterclockwise order (around the node of the spin network) for a node with + orientation, and in clockwise order for a node with − orientation, modulo cyclic permutations. Thus the ordering of the vectors is the same as the ordering of the columns of the $3j$-symbol implied by the node of the network.

This ordering is used to define a set of oriented triangles. We take the three vectors of any one of the equations (5.7) and place the base of one vector at the tip of the preceding one, to create the three edges of a triangle. In this process we parallel translate the vectors (in $\mathbb{R}^3$) but do not rotate them. The triangle is given an orientation (a definition of a normal) by taking the cross product of any two successive vectors defining the edges. For example, the normal to the 123-triangle is $\boldsymbol{J}_1 \times \boldsymbol{J}_2$, and that of the $1'4'7'$-triangle is $\boldsymbol{J}'_1 \times \boldsymbol{J}'_4$, which, in view of (5.6), is the same as $\boldsymbol{J}_1 \times \boldsymbol{J}_4$.

Next, we take the triangles and displace them so that the edge $\boldsymbol{J}_i$ of one triangle is adjacent to the edge $\boldsymbol{J}'_i$ of another triangle. In this process, the triangles are displaced but not rotated. If we do this with the six triangles defined by (5.7) in the case of the $9j$-symbol, we find that six pairs of edges can be made adjacent, as illustrated by the central six triangles of Fig. 5.2. In this "central region" six pairs of vectors $\boldsymbol{J}_i$ and $\boldsymbol{J}'_i$ are adjacent for $i = 1, 2, 5, 6, 7, 9$. There is some arbitrariness in choosing which six pairs of edges will be made adjacent. If we wish that the remaining edges $i = 3, 4, 8$ also be paired, we can duplicate three of the triangles and attach them to the periphery of the central region, as illustrated in Fig. 5.2. This amounts to a kind of "analytic continuation" of the central region.

Figure 5.2 is highly schematic. In general, the triangles are not equilateral, the surface that is formed by attaching them together is not planar, and the triangles may fold under one another.

The central region in Fig. 5.2 is a piece of an oriented surface, that is, all the normal vectors (by our convention) are pointing on the same side. In the case of the $6j$-symbol, our rules produce a closed surface (the usual tetrahedron), with normals all pointing either outward or inward (time-reversal converts one into the other). In the case of the $9j$-symbol, the surface is not closed. There is some suggestion that this surface represents a triangulation of $\mathbb{R}P^2$ but for this chapter we shall view it as living in $\mathbb{R}^3$.

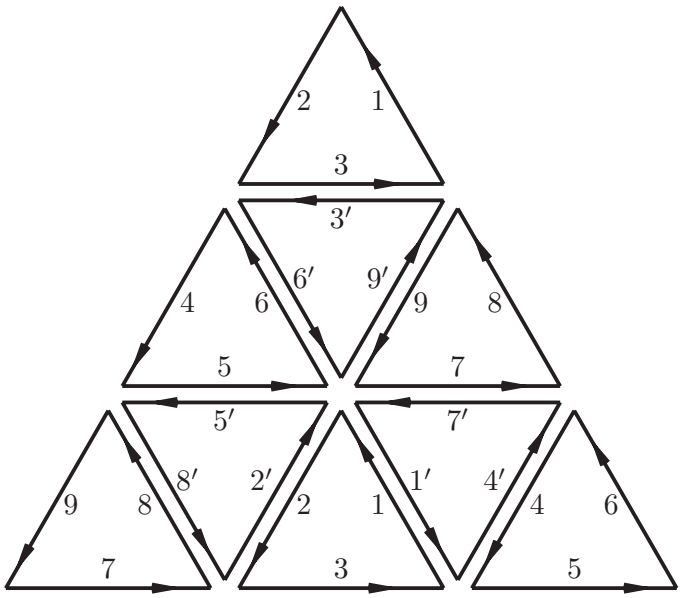


Figure 5.2: The six triangles defined by (5.7) form the "central region" of the figure, with three triangles duplicated and attached to the edges of the central region. The notation 1, $2'$, etc refers to $\boldsymbol{J}_1$, $\boldsymbol{J}'_2$, etc.

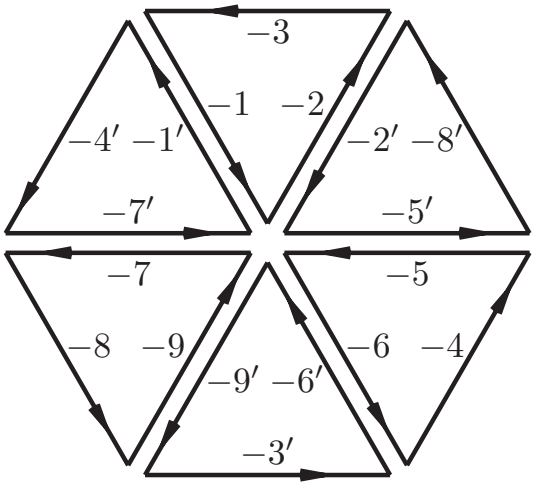


Figure 5.3: The central region of the time-reversed surface. Notation -1, $-2'$ etc refers to $-\boldsymbol{J}_1$, $-\boldsymbol{J}'_2$, etc.

Finally, we orient each edge by choosing the direction of the vector $\boldsymbol{J}_i$ (not $\boldsymbol{J}'_i$).

We will be interested in finding solutions $\{\boldsymbol{J}_i, \boldsymbol{J}'_i, i = 1, \ldots, 9\}$ of (5.5), (5.6) and (5.7), modulo overall proper rotations (in $SO(3)$). That is, although we do not rotate vectors or faces when forming our surface with oriented faces and egdes, we are allowed to rotate the whole surface once completed.

We notice that if $\{\boldsymbol{J}_i, \boldsymbol{J}'_i, i = 1, \ldots, 9\}$ is a solution of these equations, then the time-reversed set $\{-\boldsymbol{J}_i, -\boldsymbol{J}'_i, i = 1, \ldots, 9\}$ is also a solution. If we apply our rules for converting vectors into a surface, we will find in general that the time-reversed set produces a different surface (not equivalent under $SO(3)$). We apply time-reversal only to the vectors, not the rules; for example, the ordering of the time-reversed vectors is the same as the original vectors. The central six triangles of the time-reversed surface are illustrated in Fig. 5.3.

To visualize the surfaces in Figs. 5.2 and 5.3, we may imagine that the central region of Fig. 5.2 bulges out of the paper, like the northern hemisphere of a sphere (whether it does or not depends on the parameters, but this is one possibility). Then the time-reversed surface in Fig. 5.3 bulges into the paper, since spatial inversion is equivalent, modulo $SO(3)$, to reflection in a plane. Then the central region of Fig. 5.2 can be glued to the time-reversed surface in Fig. 5.3, bringing edge $\boldsymbol{J}'_3$ adjacent to edge $-\boldsymbol{J}_3$, etc, and producing a surface homeomorphic to S^2. This is the hexagonal bipyramid constructed by Ponzano and Regge [11]. The conventional normals are pointing outward in the northern hemisphere, and inward on the southern. As noted by Ponzano and Regge, this bipyramid is bisected by three planes passing through a common line, namely the "axis" of the sphere, which cut the bipyramid into three pairs of congruent tetrahedra. These correspond to the three $6j$-symbols in the representation of the $9j$-symbol as a sum over products of $6j$-symbols (see Edmonds' Eq. 6.4.3 [99]), in which the variable of summation is the common edge of the tetrahedra (the axis of the sphere).

5.4 Finding the vectors

To find a solution of (5.5), (5.6) and (5.7) we notice that all 18 vectors are determined if only four of them, $\{\boldsymbol{J}_1, \boldsymbol{J}_2, \boldsymbol{J}_4, \boldsymbol{J}_5\}$ are given. We let G be the 4×4 Gram matrix constructed out of these vectors, that is, the 4×4, real symmetric matrix of dot products of these vectors among themselves. Of the ten independent dot products, eight can be determined from the given lengths J_i, $i = 1, \ldots, 9$. That is, the diagonal elements are J_i^2, $i = 1, 2, 4, 5$, while

$$\begin{aligned} \boldsymbol{J}_1 \cdot \boldsymbol{J}_2 &= (J_3^2 - J_1^2 - J_2^2)/2, \qquad & \boldsymbol{J}_1 \cdot \boldsymbol{J}_4 &= (J_7^2 - J_1^2 - J_4^2)/2, \\ \boldsymbol{J}_2 \cdot \boldsymbol{J}_5 &= (J_8^2 - J_2^2 - J_5^2)/2, \qquad & \boldsymbol{J}_4 \cdot \boldsymbol{J}_5 &= (J_6^2 - J_4^2 - J_5^2)/2. \end{aligned} \tag{5.8}$$

The two dot products that cannot be determined from the given lengths are $u = \boldsymbol{J}_1 \cdot \boldsymbol{J}_5$ and $v = \boldsymbol{J}_2 \cdot \boldsymbol{J}_4$, which we regard as unknowns. These satisfy a linear equation

obtained by squaring $\boldsymbol{J}_9 = -\boldsymbol{J}_3 - \boldsymbol{J}_6$,

$$J_9^2 = J_3^2 + J_6^2 + 2(u + v + \boldsymbol{J}_1 \cdot \boldsymbol{J}_4 + \boldsymbol{J}_2 \cdot \boldsymbol{J}_5). \tag{5.9}$$

Another equation connecting u and v is $\det G = 0$, which holds since the four vectors lie in $\mathbb{R}^3$ and the 4-simplex defined by them is flat. This is a quartic equation in u and v, which by using (5.9) to eliminate v can be converted into a quartic equation in u alone. We write this quartic as $Q(u) = 0$. We find the roots u of this quartic, solve for v by using (5.9), whereupon all components of the Gram matrix become known (there is one Gram matrix for each root).

Ponzano and Regge [11] discussed this procedure in somewhat different language, and apparently believed that all four roots would contribute to the asymptotics of the $9j$-symbol. In fact, they do, if one wishes to work in the classically forbidden region and/or take into account tunnelling and exponentially small corrections in the neighborhood of internal caustic points (more about these below). But in the classically allowed region the asymptotics of the $9j$-symbol are dominated by the contributions from "admissible" roots, namely, those roots that produce Gram matrices that can be realized as dot products of real vectors $\boldsymbol{J}_i$. Only these correspond to real geometrical figures of the type we have described.

If a root u of $Q(u) = 0$ is complex, then it produces a complex Gram matrix that cannot be realized with real vectors, and so u is inadmissible. But a real Gram matrix can be realized as the dot products of real vectors if and only if it is positive semidefinite, so even if u is real it will still be inadmissible if G has negative eigenvalues.

We define the classically allowed region of the $9j$-symbol as the region in which $Q(u)$ has at least one admissible root. In fact, in the classically allowed region $Q(u)$ has four real roots of which two are generically admissible. We order the four real roots of $Q(u)$ in the classically allowed region in ascending order and label them by $k = 0, 1, 2, 3$. It turns out that the two admissible roots are the middle two, $k = 1, 2$, corresponding to the two terms of (5.2) with the same subscripts, $k = 1, 2$.

For a given admissible root, that is, a positive semidefinite Gram matrix, we wish to find the vectors $\boldsymbol{J}_i$, $i = 1, 2, 4, 5$. We arrange the four unknown vectors as the columns of a 3×4 matrix F, so that $G = F^T F$. To find F given G, we diagonalize G, $G = VKV^T$, where $V \in O(4)$ and K is diagonal with nonnegative diagonal entries (the eigenvalues of G). At least one of these eigenvalues must be 0; we place it last, and write $K = D^T D$ where D is a real, 3×4 diagonal matrix. Then $F = UDV^T$, where U is an arbitrary element of $O(3)$. This generates all possible sets of vectors whose dot products are realized in G; it amounts to using the singular value decomposition of F. If $U = R \in SO(3)$ then we generate a set of surfaces related by overall rotations; if $U = -R$ we generate the time-reversed set. In this way a single Gram matrix, corresponding to a single admissible root of the quartic, produces a geometry and its time-reversed image. Altogether, the two admissible roots imply the four geometries in (5.2).

This method of finding F is discussed in the context of the $6j$-symbol by Littlejohn and Yu [22], where it is also applied in the classically forbidden region. There we find complex angular momentum vectors that satisfy the required algebraic relations. This carries over to the $9j$-symbol in the classically forbidden region. In the literature on the $6j$-symbol it is common to state that a Euclidean group applies in the classically allowed region and a Lorentz group in the classically forbidden region; but for the $9j$-symbol the groups are actually $SO(3,\mathbb{R})$ and $SO(3,\mathbb{C})$.

5.5 The classically allowed region and configuration space

The classically allowed region is a subset of full dimensionality of the 9-dimensional parameter space of the $9j$-symbol, itself a convex subset of $\mathbb{R}^9$ defined by the triangle inequalities. To visualize this and other subsets of the parameter space it helps to fix seven of the j's to obtain a 2-dimensional slice. Figure 5.4 illustrates such a slice for the case

$$\left\{ \begin{array}{ccc} 129/2 & 137/2 & j_3 \\ 113/2 & 121/2 & j_6 \\ 64 & 108 & 90 \end{array} \right\}, \tag{5.10}$$

in which only j_3 and j_6 are allowed to vary. The choice of j_3 and j_6 for this purpose is not arbitrary, since these two j's are quantum numbers for a pair of commuting operators on a space of 5-valent $SU(2)$ intertwiners. They are like x and y for a wave function $\psi(x,y)$. In this analogy, we think of (j_3,j_6)-space as a "configuration space" for the $9j$-symbol and the $9j$-symbol itself as a "wave function" $\psi(j_3,j_6)$. We will mostly use the variables $J_3 = j_3 + 1/2$, $J_6 = j_6 + 1/2$ to describe this space. When thinking in classical terms, J_3 and J_6 are continuous variables (not quantized).

Figure 5.4 illustrates a convex region of the J_3-J_6 plane, bounded by straight lines and defined by the classical triangle inequalities,

$$\begin{array}{l} \max(|J_1 - J_2|, |J_6 - J_9|) \le J_3 \le \min(J_1 + J_2, J_6 + J_9) \\ \max(|J_4 - J_5|, |J_3 - J_9|) \le J_6 \le \min(J_4 + J_5, J_3 + J_9). \end{array} \tag{5.11}$$

Properly speaking, configuration space is this convex region, not the whole plane. The unshaded area inside the convex region is the classically allowed region, surrounded by the shaded classically forbidden region. The caustic curve separates the classically allowed from the classically forbidden regions; it has kinks (discontinuities in slope) at points B, and is tangent to the boundary of the convex region at several points. Other features of this figure are explained below.

Given a point (J_3, J_6) of the classically allowed region, the procedure described in sections 5.3 and 5.4 produces a quartic polynomial $Q(u)$ whose two middle roots $k = 1, 2$ are admissible. These can be thought of as specifying a two-branched "root

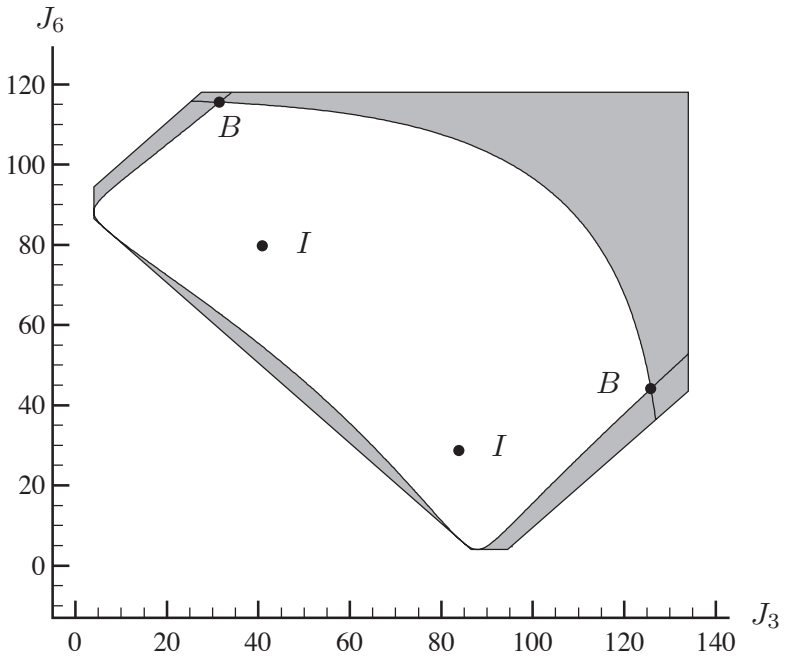


Figure 5.4: The convex region of the J_3-J_6 plane is the configuration space of the $9j$-symbol. The shaded area is the classically forbidden region, and the unshaded, the classically allowed. Points I are internal caustics, two of the flat configurations; points B are the other two flat configurations, lying on the boundary curve.

surface" that sits over the classically allowed region. The two middle roots coalesce as we approach the caustic curve, and become (inadmissible) complex conjugates as we move beyond. Thus, the two root surfaces can be thought of as being glued together on the caustic curve.

Corresponding to each root there are two geometries modulo $SO(3)$, related by time-reversal, so there is a two-fold "geometry surface" sitting above each root surface, or four geometry surfaces sitting above the classically allowed region. These four geometry surfaces are actually branches of the projection of an invariant 2-torus onto configuration space, and correspond to the four exponential terms in (5.2). This 2-torus sits in the phase space of the $9j$-symbol, a 4-dimensional, compact symplectic manifold.

This symplectic manifold is only one of several phase spaces that describe the classical mechanics of the $9j$-symbol, but all the others have higher dimensionality so we call this one the "phase space of minimum dimensionality." It is one of the symplectic manifolds discovered by Kapovich and Millson [110]. Consequently this phase space is isomorphic to the space of shapes $\mathcal{P}_5$, although in the present context the angular momentum vectors $\boldsymbol{J}_i$ are interpreted as edge vectors rather than area

vectors. Its analog in the case of the $6j$-symbol is a spherical phase space isomorphic to the one studied at length in the last chapter. This space has also been studied by Charles [21] and by Littlejohn and Yu [22]. The phase space of minimum dimensionality is related to other phase spaces for the $9j$-symbol by a combination of symplectic reduction [57] and the elimination of constraints. We have found it useful to employ all these spaces in our work on the $9j$-symbol.

5.6 The amplitude and caustics

The amplitudes of semiclassical approximations are notorious for the computational difficulties they cause. For example, several authors have resorted to computer algebra and/or numerical experimentation to check the amplitude determinant in the Ponzano-Regge formula. Actually, this amplitude (due originally to Wigner [138]) is given by a single Poisson bracket between intermediate angular momenta (see [37], and, in more detail, [22]), which can be evaluated in a single line of algebra. More generally, semiclassical amplitudes are easily found in terms of matrices of Poisson brackets, as discussed in Chapter 2.

In the case of the $9j$-symbol we define

$$V_{ijk} = \boldsymbol{J}_i \cdot (\boldsymbol{J}_j \times \boldsymbol{J}_k), \tag{5.12}$$

which is six times the signed volume of the tetrahedron specified by edges i, j, k (it is the volume of the corresponding parallelepiped). Then the amplitudes A_1, A_2 in (5.2) are given by

$$A = \frac{1}{4\pi\sqrt{|\det D|}}, \tag{5.13}$$

where

$$D = \begin{pmatrix} V_{124} & V_{215} \\ V_{451} & V_{542} \end{pmatrix}. \tag{5.14}$$

The subscripts 1,2 are omitted on A in (5.13) because the same formula applies for both terms in (5.2), but $A_1 \neq A_2$ in general because the formula is evaluated on two different geometries (associated with the two admissible roots). The quantitity $\det D$ is even under time-reversal, so the same amplitude applies to both a geometry and its time-reversed image.

The volumes in matrix D are Poisson brackets of intermediate angular momenta in a recoupling scheme for the $9j$-symbol, which are most easily evaluated in the phase space of minimum dimensionality. We omit details; suffice it to say for now that the derivation of the matrix (5.14) in terms of Poisson brackets and thence the amplitude is extremely easy.

We define the caustic set as the subset of the $9j$-parameter space where $\det D = 0$. Its intersection with the 2-dimensional slice seen in Fig. 5.4 consists of the union of

the caustic curve (the curve separating the classically allowed from the classically forbidden region) with the two points marked I. In addition, the caustic set includes the continuation of the caustic curves from points B into the classically forbidden region. The points I are "internal" caustics, that is, internal to the classically allowed region. While the caustic curve has codimension 1, the internal caustics have codimension 2.

The quantity $\det D$ is nonzero away from the caustics. It turns out that the sign of $\det D$ distinguishes the two root surfaces, with $\det D > 0$ on root surface 1 and $\det D < 0$ on root surface 2.

The caustics of the $6j$-symbol occur at the flat configurations (flat tetrahedra), as appreciated by Ponzano and Regge [11] and Schulten and Gordon [139, 140]. The caustics of the $9j$-symbol, however, are not in general flat, that is, $\det D = 0$ does not imply that the configuration is flat. The flat configurations of the $9j$-symbol, however, do lie on the caustic set. In a given J_3-J_6 slice, there are precisely four flat configurations. In the example of Fig. 5.4, these are marked B and I. The points B are flat configurations lying on the boundary of the classically allowed region (the caustic curve), while points I are internal flat configurations. As we vary the seven j's that are fixed in Fig. 5.4, the number of flat configurations on the boundary varies from 2 to 4; those not on the boundary are internal.

In the usual manner of semiclassical approximations, (5.2) breaks down in a neighborhood of the caustic set (it diverges exactly at the caustic), and must be replaced by a diffraction function associated with a catastrophe [133]. In the case of the $6j$-symbol, the only catastrophe that occurs is the fold, yielding an Airy function as the semiclassical approximation, as noted by [11, 139, 140]. This is the normal situation for systems of one degree of freedom. The $9j$-symbol, however, possesses two degrees of freedom, and other types of catastrophes occur. The fold catastrophe applies at most points along the caustic curve, where the $9j$-symbol is approximated by an Airy function; but at flat configurations there is an umbilic catastrophe, hyperbolic for those (B) falling on the boundary (caustic) curve and elliptic for the internal caustics (I). See Trinkaus and Drepper [141] for illustrations of the associated diffraction functions. The umbilic catastrophes are generic in systems of three degrees of freedom but occur in the $9j$-symbol (with only two) because of time-reversal symmetry. However, only sections of the full three-dimensional umbilic wave forms appear [133]. The cusp catastrophe, which can be expected in generic systems of two degrees of freedom, does not occur in the classically allowed region of the $9j$-symbol.

Caustics are associated with the coalescence of branches of the projection of a Lagrangian manifold in phase space onto configuration space. In the case of the $9j$-symbol, the Lagrangian manifold is the invariant 2-torus mentioned in section 5.5. Along the boundary of the classically allowed region, the two admissible roots coalesce, which means that the four geometries merge into two. At most points on the boundary curve, the two remaining geometries are not equal, but are related by time-reversal. At such points we have a fold catastrophe, and the $9j$-symbol is approximated by an Airy function (modulated by a cosine term). At points B, however, the two

geometries related by time-reversal merge into a single flat configuration, producing the hyperbolic umbilic catastrophe.

At internal caustic points I the geometry and its time reversed image for one of the two admissible roots coalesce to produce a flat configuration. The two geometries of the other root surface, however, do not coalesce. Thus at internal caustics I there are three geometries. Only the flat configuration associated with one of the roots produces the elliptic umbilic catastrophe; thus, only one of the two terms in (5.2) is replaced by the elliptic umbilic diffraction function, while the other remains as shown in (5.2). The $9j$ symbol is a linear combination of these two terms, but the elliptic umbilic diffraction function dominates when the scaling factor k is large.

The caustics have a certain size, that is, a distance around the caustic set over which diffraction functions must be used instead of (5.2). This distance Δj scales as $k^{1/3}$ for all three catastrophe types (fold and elliptic and hyperbolic umbilic) discussed here.

In the neighborhood of fold catastrophes the wave function scales as $k^{-17/6}$, that is, $k^{1/6}$ higher than the k^{-3} of the two terms in (5.2). In the neighborhood of umbilic catastrophes the scaling is $k^{-8/3}$, that is, with another factor of $k^{1/6}$. For large values of k the $9j$-symbol is largest near the points I, B. Figure 5.5 displays the exact values of the $9j$-symbol for the parameters of Eq. (5.10) and corresponds to the configuration space pictured in Figure 5.4. All three types of caustic points (fold, elliptic and hyperbolic umbilic) are visible in Fig. 5.5 due to the wavefunction enhancements at these points. The elliptic umbilic at the point I near $(J_3 = 40, J_6 = 80)$ is particularly noticeable.

Linear combinations with different scaling behaviors have been observed by Barrett and Steele and by Freidel and Louapre in their studies of the $10j$-symbol [18, 19]. It seems that the $9j$-symbol is the simplest spin network in which this phenomenon occurs.

5.7 The phase

The phases S_1 and S_2 in (5.2) each have the form

$$S = \sum_{i=1}^{9} J_i \theta_i, \tag{5.15}$$

where θ_i is the angle between normals of adjacent faces of the geometrical figure. This of course is similar to the Ponzano-Regge formula, but the $6j$-tetrahedron is convex and all dihedral angles can be taken in the interval $[0, \pi]$. The dihedral angles for the $9j$-symbol, on the other hand, must be allowed to lie in a full 2π interval, as explained momentarily. The subscripts 1,2 are omitted on S in (5.15) because the same formula applies to both terms in (5.2). The formula must be evaluated, however,

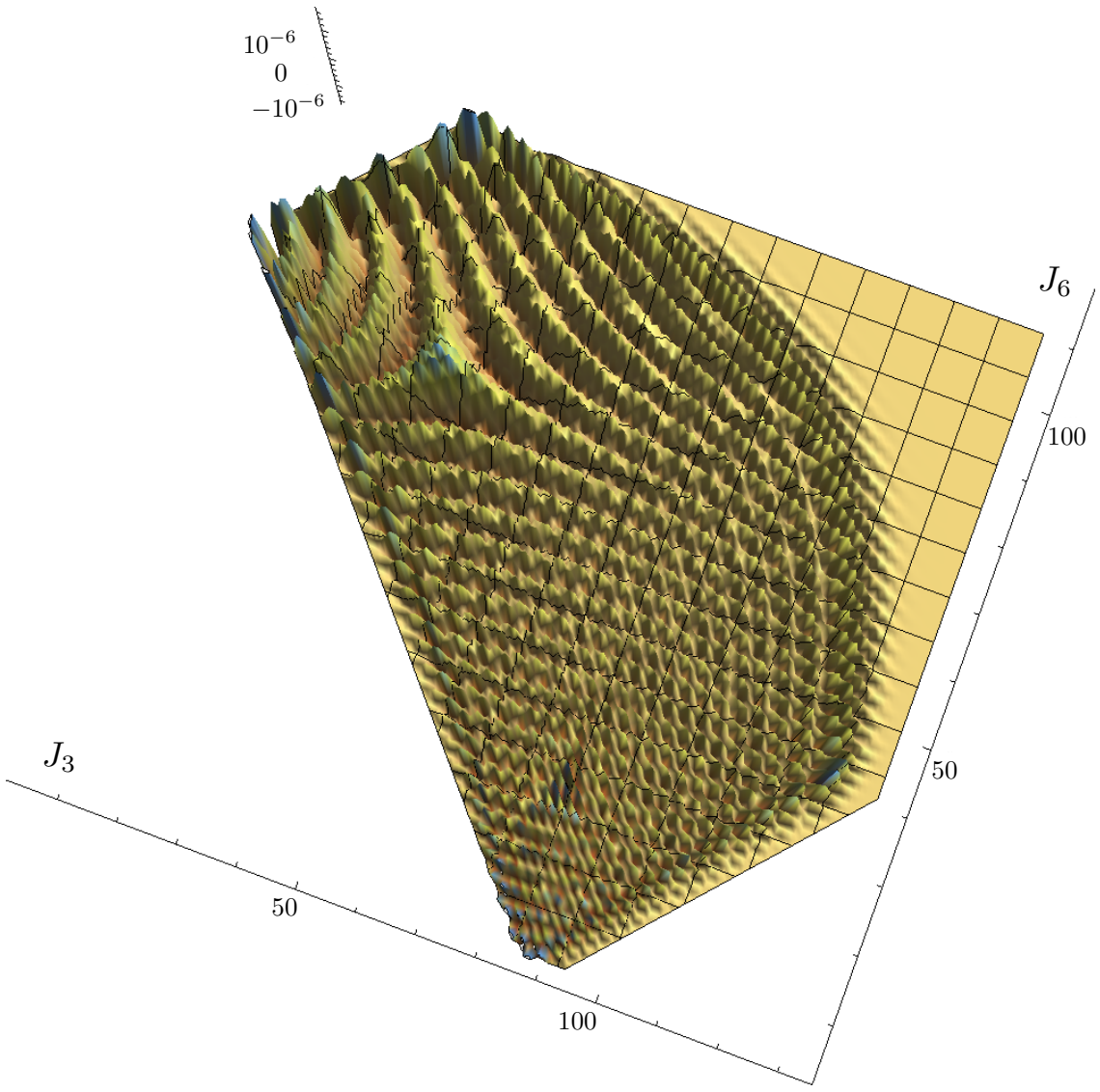


Figure 5.5: The exact wavefunction of the $9j$-symbol. Like Fig. 5.4 the parameters of the $9j$-symbol are given by Eq. (5.10). All three types of caustic points (fold, elliptic and hyperbolic umbilic) are visible due to the wavefunction enhancements at these points.

on two different geometries, so S_1 and S_2 are not equal. In addition, the angles θ_i lie in different intervals for the two geometries.

Each edge i of the geometrial figure is adjacent to two faces, for example, edge 4 in Fig. 5.2 is adjacent to faces $1'4'7'$ and 456. One face adjacent to edge i contains vector $\boldsymbol{J}_i$, and the other $\boldsymbol{J}'_i$. Let the two normals of these two faces, according to the conventions given above, be $\hat{\mathbf{n}}$ and $\hat{\mathbf{n}}'$. Then we define θ_i as the angle such that

$$R(\hat{\boldsymbol{\jmath}}, \theta_i)\hat{\mathbf{n}} = \hat{\mathbf{n}}', \tag{5.16}$$

where $\hat{\boldsymbol{\jmath}}$ is the unit vector along $\boldsymbol{J}$, specifying the axis of a rotation R by angle θ_i using the right-hand rule. In the Ponzano-Regge formula one can compute the dihedral angle from its cosine, but for the $9j$ one must also use the sine of the angle. That is, (5.16) is equivalent to

$$\hat{\mathbf{n}}' = \cos\theta_i\, \hat{\mathbf{n}} + \sin\theta_i\, \hat{\boldsymbol{\jmath}} \times \hat{\mathbf{n}}. \tag{5.17}$$

This determines θ_i to within an additive integer multiple of 2π. We add the further requirement that for the geometries associated with the first root (the cosine term in (5.2)), $-\pi \le \theta_i < +\pi$, while for the second root (the sine term in (5.2)), $0 \le \theta_i < 2\pi$. These ranges for the angle θ_i are chosen because they give a continuous branch for the angle over the two root surfaces. It turns out that θ_i never crosses $\pm\pi$ on the surface for root 1, and it never crosses 0 or 2π on the surface for root 2.

The rules given in sectionss 5.3 and 5.4 for converting vectors into surfaces with oriented edges and triangles are an essential part of the definition of the dihedral angles θ_i. It is of interest to see how the angles change when a set of vectors or the associated geometry is subjected to some symmetry.

Under time-reversal, the orientation of all triangles reverses, that is, the normal vectors stay the same but the vectors defining the edges are inverted. This means that the angles θ_i go into $-\theta_i$ on root surface 1, while they go into $2\pi - \theta_i$ on root surface 2 (both changes guarantee that the angles remain within their respective ranges). Thus S goes into $-S$ on root surface 1 and

$$S \to -S + 2\pi\nu + 9\pi \tag{5.18}$$

on root surface 2, where ν is the integer,

$$\nu = \sum_{i=1}^{9} j_i. \tag{5.19}$$

These guarantee that $\cos S_1$ and $\sin S_2$ are invariant under time-reversal. Since the same applies to the amplitudes A_1 and A_2, one can choose either a geometry or its time-reversed image, for each root, when evaluating (5.2).

This completes the definition and geometrical interpretation of all the notation used in (5.2).

5.8 Symmetries of the $9j$-symbol

The formula (5.2) transforms correctly under the symmetries of the $9j$-symbol ([101], Sec. 10.4), which state that the $9j$-symbol suffers a phase change of $(-1)^\nu$ under odd permutations of rows or columns or under transposition. Consider, for example, the swapping of the first two columns, and let P be the permutation of indices, so that $P1 = 2$, $P2 = 1$, $P3 = 3$, etc. This maps an old set of nine j's into a new set, and old quartic $Q(u) = 0$ into a new one, etc. We find that the u root of the old quartic becomes the v root of the new one, which amounts to saying that the root 1 surface of the old geometry is mapped into the root 2 surface of the new one, and vice versa. Also, the orientations of the three unprimed triangles reverse, but not those of the primed ones, causing all nine dihedral angles to be incremented or decremented by π (depending on the range). If we let θ_i be the original angles and $\tilde{\theta}_i$ the new ones, then when θ_i is on root surface 1 we find $\tilde{\theta}_{Pi} = \theta_i + \pi$, which means that the new angle is in the right range since it is on root surface 2. Similarly, when θ_i is on root surface 2 then $\tilde{\theta}_{Pi} = \theta_i - \pi$, which is in the right range since $\tilde{\theta}_{Pi}$ is on root surface 1. As a result, when the original geometry is on root surface 1, we have

$$\sum_{i=1}^{9} J_i \tilde{\theta}_i = \sum_{i=1}^{9} J_i \theta_i + \nu\pi + \frac{9\pi}{2}, \tag{5.20}$$

so that $\sin \tilde{S}_2 = (-1)^\nu \cos S_1$, while if the original geometry is on root surface 2, we have

$$\sum_{i=1}^{9} J_i \tilde{\theta}_i = \sum_{i=1}^{9} J_i \theta_i - \nu\pi - \frac{9\pi}{2}, \tag{5.21}$$

so that $\cos \tilde{S}_1 = (-1)^\nu \sin S_2$. The sine and cosine terms in (5.2) swap under column swap, and the result acquires an overall phase of $(-1)^\nu$, as required. The specified ranges on the dihedral angles on the two root surfaces are necessary for this to work out.

5.9 Conclusions and extensions

It is easy to derive the expression (5.15) by the method of Roberts [17], which involves rotating faces by an angle of π about their normals, and edges by an angle of π about a normal to them. The phase (5.15) (times 2) is then an action integral along one Lagrangian manifold and back along another (the analogs of the A- and B-manifolds of [37]). Similar expressions apply to any spin network of any complexity. But the contours chosen for the integration are not unique, in that one can add any multiples of quantized loops on the two manifolds. These modify both the actions and the Maslov indices, and amount to changing the choice of branch for the angles θ_i, that is, adding an integer multiple of 2π to these angles. This does not leave the

trigonometric functions in (5.2) invariant because the angles are multiplied by the J_i, which may be half-integers. The result is that the phase of the approximation to the $9j$-symbol depends on the contours. A more serious worry is that the contours, that is, the branches for the θ_i, may change as we move around in the parameter space of the $9j$-symbol. This would amount to crossing a branch cut for the angles θ_i (and there are different branch cuts for different angles). In addition, as we move around in parameter space we can make any two adjacent faces rotate relative to one another around their common edge as many times as we want. Although the phases in question are "only" powers of -1, straightening out this issue was by far the hardest part of this work. In the end we realized that the ranges $[-\pi, +\pi)$ on root surface 1 and $[0, 2\pi)$ on root surface 2 guarantee that there are no branch cuts and hence no discontinuities. The ranges specified for the angles θ_i give us in effect a global, smooth definition of contours for carrying out action integrals.

We present several numerical comparisons of (5.2) with the exact $9j$-symbol. In Fig. 5.6 the approximation (5.2) (smooth curve) may be compared to the exact $9j$-symbol (sticks) as a function of j_3 for fixed values of the other j's. The range chosen lies inside the classically allowed region, far from a caustic. Fig. 5.7 shows the comparison in a range that crosses a fold catastrophe, and Fig. 5.8 shows the comparison in an interval that passes near a hyperbolic umbilic catastrophe (the upper point I in Fig. 5.4). The approximation (5.2) is too large near the point I.

Varshalovich *et al* [101] present an asymptotic approximation for the $9j$-symbol without citation (their Eq. (10.7.1)), which is different from our formula (5.2). In Figure 5.9 we compare the exact $9j$-symbol with the formula of Varshalovich *et al* and with our formula (5.2), for the values

$$\left\{ \begin{array}{ccc} 32 & 34 & j_3 \\ 28 & 61/2 & 81/2 \\ 26 & 73/2 & 91/2 \end{array} \right\}. \tag{5.22}$$

The formula of Varshalovich *et al* vanishes at many places inside the classically allowed region, as we have defined it, so it takes some searching to find an interval where both their formula and ours give nonzero results. On the basis of such comparisons, we believe that the formula of Varshalovich *et al* is an asymptotic result in a different sense than ours, or else it is incorrect.

The two terms in (5.2) have different trigonometric functions (sine and cosine) because there is a relative Maslov index of 2 between the two root surfaces. The relative Maslov index between a geometry and its time-reversed image is 0, a somewhat surprising result because in mechanical systems and in the $6j$-symbol the Maslov index between a branch or geometry and its time-reversed image is 1.

When an interior caustic occurs on a root surface, the two geometries that sit above it form a double cover, in the manner of the Riemann sheet for the square root function. The internal caustic point I is a branch point for the cover. Geometries

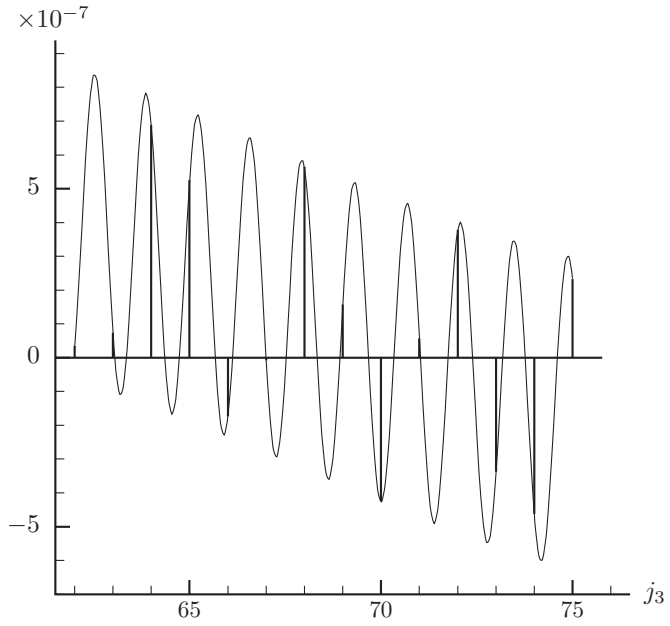


Figure 5.6: Comparison of exact $9j$-symbol (vertical sticks) with approximation (5.2), away from a caustic. Values used are those in (5.10), with $j_6 = 50$.

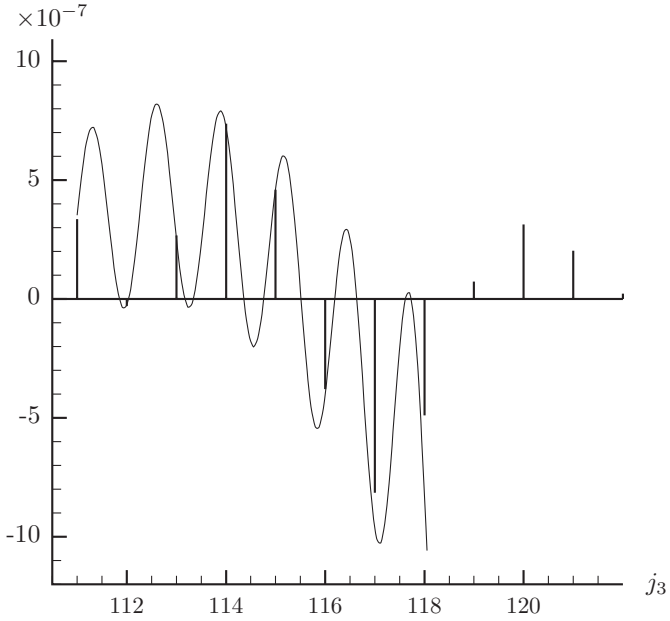


Figure 5.7: Like Fig. 5.6, but an interval that spans a fold catastrophe (with $j_6 = 60$). The approximation (5.2) is discontinued at the caustic, the exact values are continued into the classically forbidden region.

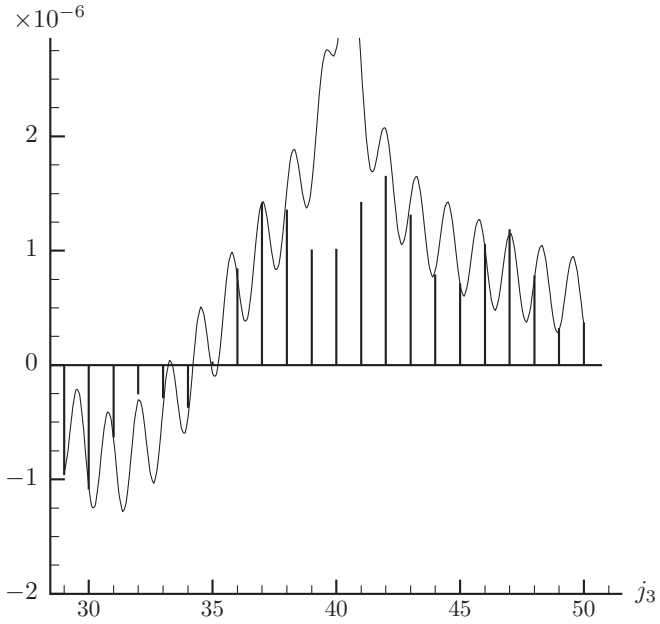


Figure 5.8: Like Fig. 5.7, but passing near an ellipitc umbilic catastrophe (with $j_6 = 79$).

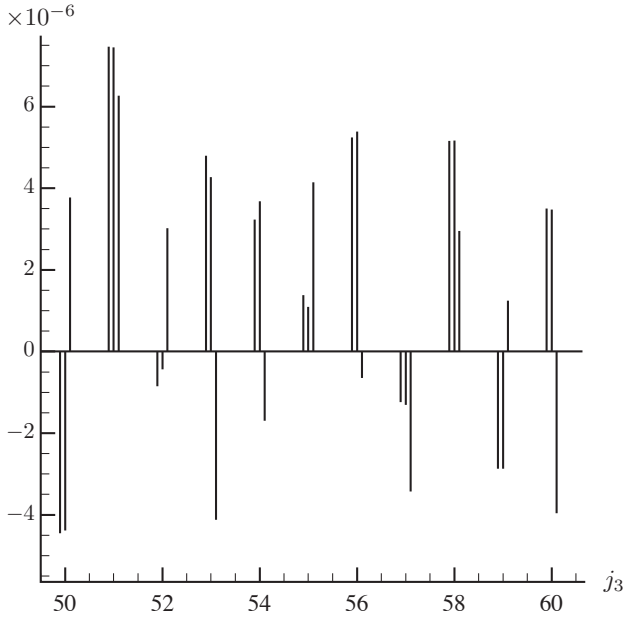


Figure 5.9: Near each quantized value of j_3, there are three lines. The other eight j's are given by (5.22). The left line is the exact $9j$-symbol, the middle line is approximation (5.2), and the right is formula (10.7.1) of Varshalovich *et al* [101].

transform continuously into their time-reversed images as we go around the point I, without crossing a caustic.

Several studies of the asymptotics of spin networks have started with an integral representation of the network, to which the stationary phase approximation is applied. Roberts represented the $6j$-symbol as a scalar product in a certain Hilbert space, which was put into the coherent state representation, whereupon the integral was evaluated by the stationary phase approximation [17]. Coherent states have played a prominent role in many recent semiclassical studies. Our approach has been to work as much as possible in a representation-independent manner. For example, the stationary phase points are seen as intersections of Lagrangian manifolds. Some of the basics of this approach were presented in [37]. We have not specifically used the coherent state or any other representation.

Some aspects of this calculation carry through in an obvious way to higher spin networks, while for others nontrivial generalizations seem to be required. One extension necessary for the study of higher $3nj$-symbols is the admissible root problem. The fact that the geometries corresponding to the $9j$-symbol were completely determined by the quartic $Q(u)$ was key. We have also found a practical solution to the root determination problem for higher $3nj$-symbols. For definiteness let us focus on the $12j$-symbol but the approach should also apply to the higher spin networks. In the case of the $12j$-symbol all 24 vectors are determined if only five of them are given, $\{\boldsymbol{J}_1, \ldots, \boldsymbol{J}_5\}$. Once again we require that these five vectors lie in $\mathbb{R}^3$ and so the determinant of the 5×5 Gram matrix vanishes $\det G_5 = 0$. It is also necessary for all 4×4 principal minors to vanish, this is related to Sylvester's criterion [142]. The conditions imposed by the coupling scheme once more gives rise to linear relations between the dot products and these couple the principle minors in such a way that the vanishing of two principle minors implies the vanishing of the others (and of $\det G_5$). The admissible roots of the $12j$-symbol are completely determined by this system of two coupled quartics and the linear relations of the coupling scheme. Following the recommendations of P. Rostalski and C. Vinzant we have been able to solve this system numerically using polynomial homotopy continuation (see Sommese and Wampler for an introduction [143]). The number of admissible roots and the caustic set of the $12j$-symbol can be studied using this technique. A very rich structure emerges and we hope to report at length on this work in the future.

This short introduction has barely scratched the surface of the study of the $9j$-symbol. We have collected a number of results not reported here: a geometrical understanding of the tangencies of the caustic to the triangle allowed boarders, an explicit calculation of the Maslov indices using the techniques introduced in Chapter 2 and a proof that the number of internal caustics can be 0, 1 or 2 but not 3 or 4. There are also more speculative lines to pursue: The Lagrangian manifold of the $9j$-symbol is a torus in the lowest dimensional phase space. The four caustic points I and B, corresponding to four flat configurations of the vector diagram, act as marked points on the torus. This description is remarkable close to the setup of elliptic function

theory and we conjecture that elliptic functions will also play an important role in future studies of the $9j$-symbol.

Certainly a deeper understanding of the $9j$ results is necessary for a full understanding of the asymptotics of higher spin networks. Due to the important role played by the $15j$-symbol in the spinfoam models of loop gravity this will be an exciting avenue for researchers interested in quantum gravity and semiclassics to pursue.

Chapter 6

Conclusion

In this dissertation we have surveyed a number of applications of semiclassics to spin networks and quantum gravity. One of the most exciting features of this work is the large number of avenues for future research that it suggests, we discuss some of these here.

Chapter 4 provided a thorough investigation of the volume of a quantum tetrahedron, which we have argued is associated to the volume of a 4-valent node of a spin network in quantum gravity. Several questions arise in attempts to extend this work to polyhedra with more faces and higher valent nodes of a spin network. We have discovered a solution to the Minkowski reconstruction problem for triangular prisms and in principle our technique may be extended to polyhedra with more faces. However, this approach had an important enumerative step and as the number of faces of the polyhedron grows there is an exponential explosion in the number of adjacency classes. Is there a more incisive and abstract way of capturing this procedure? If not, how far can we push the algorithm before it becomes impractical? Already in the case of six faces there appears a new phenomenon. Instead of there being a single dominant class (the triangular prism) with a single degenerate case (the quadrilateral pyramid) there are two dominant classes (cuboids and pentagonal wedges) and many degenerate cases.

The most interesting questions relate to the volume evolution of the the triangular prism. Can the adjacency class of the prism change along the orbits of the volume? Is this evolution chaotic? And if it is, what does this imply about the volume spectrum of loop quantum gravity? We have briefly mentioned that symplectic integrators provide practical techniques for investigating these questions. The highly non-linear structure of the classical volume of the prism suggests that the dynamics will be chaotic. If periodic orbits can be identified within this dynamics then the density of states could be investigated using Gutzwiller's trace formula [132, 63]. What aspects of quantum gravity could be investigated using this density of states?

In Chapter 5 we have laid the foundations for systematic investigation of the asymptotics of higher 3nj-symbols. Our work in [37, 24] has shown that there are

simple formulas that can be used to calculate the amplitudes of these symbols and the phases are well controlled using Roberts' methods [17]. Here we have discussed the fact that polynomial homotopy continuation can be used to find the admissible roots which give rise to the geometries associated to a spin network. All of this points to the fact that the asymptotics of higher $3nj$-symbols should be analytically accessible. Furthermore these tools show that the spin networks can be turned into practical computational objects and we have now to look for quantitative applications and physical predictions that can be made with these objects.

Our work on the higher $3nj$-symbols has focused on the "edge length picture", that is, on the asymptotics as they relate to $SU(2)$. More interesting for applications to quantum gravity is the area picture of the Minkowski theorem and the relationships between the $9j$-symbol (determined by 5 edge vectors) and the triangular prism (determined by 5 area vectors) have not even begun to be studied. The loop gravity literature has primarily focused on the spin networks of other groups, groups more closely related to the geometry of 4-space ($SO(4)$ and $SL(2,\mathbb{C})$), and there is a much to do in connecting their work to ours.

A subject that we have not been able to discuss at all is the asymptotics of the q-deformed symbols. The q-deformations have proved to be immensely interesting mathematically and there are indications that they relate to the cosmological constant in loop gravity. The asymptotics of the q-deformed $6j$-symbol have been studied by Woodward and Taylor [26, 27] and they find that these asymptotics are related to the geometry of curved tetrahedra (hyperbolic for real q and spherical for q a root of unity). The methods of Woodward and Taylor are similar to those used by Schulten and Gordan [139, 140] in their early studies of the usual $6j$-symbol. These results cry out for a geometrical interpretation and proof. More recent work of Kapovitch, Millson and Treloar may provide a framework for just such a proof [144]. We also wonder whether there is a generalization of the Minkowski theorem to spaces of constant curvature. And if so, what are the asymptotics of the volume operator in this context?

Our derivation of an asymptotic formula for the $9j$-symbol has already been put to good use by Littlejohn and Yu and Yu in [145, 146, 147]. They have derived new asymptotics for $3nj$-symbols containing a mixture of large and small j's. Once more these asymptotics seem like they will provide a useful tool in the study of loop quantum gravity.

At several junctions we have noted relationships between the $3nj$-symbols and objects in atomic and molecular physics. An important example of this that we have not yet mentioned is that the $9j$-symbol is precisely the matrix that takes you from an LS-coupling scheme to a jj-coupling scheme. We have not even begun to think about the application of our asymptotic formulae for the volume operator and the $9j$-symbol to this context and it seems to be a rich area for them.

Were there more time we would add a chapter to this dissertation providing a smoother bridge between its two parts, Foundations and Applications. This chapter

would discuss the Regge calculus at greater length. The Regge calculus is well studied but there are still many directions in which it needs to be developed. For example, we have not been able to find a thorough development of the Regge calculus in an analogous manner to the first order formulation of general relativity (for beginnings see [148]). Many aspects of general relativity, such as its canonical structure and action principles, simplify in a first order formulation. This is also true of Regge calculus, see [149]. There are two interesting areas of Regge calculus that we can glimpse the outlines of but which we have not yet seen formally treated, the first is a notion of piecewise linear parallel translation and the second is a notion of Legendre transform. Both of these relate to a first order formulation.

Early in the study of parallel transport Cartan introduced a connection as the notion of a rolling frame (for an introduction see [150]). Take a model constant curvature space, e.g. the tangent space of the manifold of interest, and roll it along the curve to be parallel translated along in the manifold. For example, picture a plane tangent to a sphere being rolled along a curve in the sphere so that the point of tangency is always on the curve. This tells you how to translate the vectors of the tangent space. The exact same idea can be used for the simplices of the Regge calculus. Consider a tetrahedron and the tangent plane to one of its faces. Once again this tangent plane can be rolled around the tetrahedron and as it transitions from one face to the next it rolls about the edge connecting the two faces in an amount equal to the external dihedral angle between them. The "connection" on the tetrahedron is determined by its external dihedral angles. This is remarkably reminiscent of the Regge calculus and of the Ponzano-Regge formula for the $6j$-symbol. This exact procedure is also the foundation for defining projective duality in the context of algebraic geometry. In fact, this is a generalization of the geometrical description of the Legendre transform [151]. Is there a precise sense in which the angles and the edge lengths appearing in the $3nj$ asymptotics are conjugate variables? What is the symplectic manifold for which the Ponzano-Regge phase $S_{PR} = \sum_i J_i \theta_i$ acts as a generating function?

It was a revelation for us that the Ponzano-Regge phase of the $6j$-symbol is a discretization of the *boundary* action of general relativity, an analog of the Gibbons-Hawking-York term [16]. The similarity between this phase and the bulk Regge action are remarkable. Following Padmanabhan's work in classical general relativity [152], we have realized that there is a precise mathematical relationship between these bulk and boundary actions and along with him we call this a holographic relationship.[1] Padhmanabhan has shown that the holgraphic correspondence of bulk and boundary actions in general relativity can be used to provide a classical thermodynamic picture of black holes. Can a similar picture be developed in the Regge calculus? This would be immensely useful for comparing to first principle derivations of black hole entropy

[1]Note that this is a different, although not unrelated, sense of holographic than is used in the string theory literature.

in loop gravity and other discrete approaches to gravity.

In 1751 Carl Linnaeus, the great botanist, physician and zoologist, designed a flower garden that measured time. If the orange hue of a scarlet pimpernel was visible then it was time to breakfast (8am) and if a hawkweed closed its petals it was time for an afternoon nap (3-4pm). The pace of life has changed since then but time is still written in all of our activities. The unmistakable finger print of a city is the tempo at which its inhabitants walk. From this tempo you can predict the city's average income, number of restaurants, number of colleges and most importantly its size [153, 154]. Throughout this dissertation we have ignored time, approximated it away for the convenience of our calculations. Does general relativity's unification of space and time together with the quantization of space studied here imply that the flow of time is discrete?

Bibliography

[1] C. Rovelli, *Quantum Gravity*. Cambridge University Press, 2004.

[2] T. Thiemann, *Modern Canoncial Quantum General Relativity*. Cambridge University Press, 2007.

[3] A. Ashtekar, "An introduction to loop quantum gravity through cosmology," *Nuov. Cim. B* **122** (2007) 135–155, arXiv:gr-qc/0702030.

[4] A. Perez, "Introduction to Loop Quantum Gravity and Spin Foams," *ArXiv e-prints* (2004) arXiv:gr-qc/0409061.

[5] S. Mercuri, "Introduction to Loop Quantum Gravity," *ArXiv e-prints* (2010) 1001.1330.

[6] P. Doná and S. Speziale, "Introductory lectures to loop quantum gravity," *ArXiv e-prints* (2010) 1007.0402.

[7] C. Rovelli, "Zakopane lectures on loop gravity," *ArXiv e-prints* (2011) 1102.3660.

[8] L. Freidel and S. Speziale, "Twisted geometries: A geometric parametrization of SU(2) phase space," *Phys. Rev. D* **82** (2010), no. 8, 084040–1–16, 1001.2748.

[9] L. Freidel and S. Speziale, "From wistors to twisted geometries," *Phys. Rev. D* **82** (2010), no. 8, 084041–1–5, 1006.0199.

[10] T. Regge, "General relativity without coordinates," *Nuov. Cim.* **19** (1961) 558–571.

[11] G. Ponzano and T. Regge, "Semiclassical limit of Racah coefficients," in *Spectroscopy and Group Theoretical Methods in Physics*, F. B. et al, ed., p. 1. Amsterdam: North-Holland, 1968.

[12] L. M. Blumenthal, "A budget of curiosa metrica," *Amer. Math. Monthly* **66** (1959) 453–460.

[13] E. Anderson, "The Problem of Time in Quantum Gravity," *ArXiv e-prints* (2010) 1009.2157.

[14] R. Oeckl, "General boundary quantum field theory: Foundations and probability interpretation," *Adv. Theor. Math. Phys.* **12** (2008) 319–352, arXiv:hep-th/0509122.

[15] E. Bianchi, "Loop Quantum Gravity." http://wwnpqft.inln.cnrs.fr/pdf/Bianchi.pdf, 2010. Lectures at the Winter Workshop on Non-Perturbative QFT Sophia-Antipolis.

[16] J. Hartle and R. Sorkin, "Boundary terms in the action for the Regge calculus," *Gen. Rel. Grav.* **13** (1981) 541–549. 10.1007/BF00757240.

[17] J. Roberts, "Classical 6j-symbols and the tetrahedron," *Geom. Top.* **3** (1999) 21–66, arXiv:math-ph/9812013.

[18] J. W. Barrett and C. M. Steele, "Asymptotics of relativistic spin networks," *Class. Quant. Grav.* **20** (2003), no. 7, 1341.

[19] L. Freidel and D. Louapre, "Asymptotics of 6 j and 10 j symbols," *Class. Quant. Grav.* **20** (2003), no. 7, 1267.

[20] R. Gurau, "The Ponzano-Regge asymptotic of the $6j$ symbol: an elementary proof," *Ann. Henri Poincare* **9** (2008) 1413–1424.

[21] L. Charles, "On the quantization of polygon spaces," *ArXiv e-prints* (2008) 0806.1585.

[22] R. G. Littlejohn and L. Yu, "Uniform semiclassical approximation for the Wigner $6j$-symbol in terms of rotation matrices," *J. Phys. Chem. A.* **113** (2009) 14904–14922.

[23] M. Dupuis and E. R. Livine, "Pushing the asymptotics of the $6j$-symbol further," *Phys. Rev. D* **80** (2009) 024035.

[24] V. Aquilanti, H. M. Haggard, A. Hedeman, N. Jeevanjee, R. G. Littlejohn, and L. Yu, "Semiclassical Mechanics of the Wigner $6j$-Symbol," *ArXiv e-prints* (2010) 1009.2811.

[25] S. Mizoguchi and T. Tada, "Three-dimensional gravity from the Turaev-Viro invariant," *Phys. Rev. Lett.* **68** (1992) 1795–1798.

[26] Y. U. Taylor and C. T. Woodward, "Spherical Tetrahedra and Invariants of 3-manifolds," *ArXiv e-prints* (2004) arXiv:math/0406228.

[27] Y. U. Taylor and C. T. Woodward, "6j symbols for non-Euclidean tetrahedra," *Sel. Math. New* **11** (2005) 539–571.

[28] J. W. Barrett and R. M. Williams, "The asymptotics of an amplitude for the 4-simplex," *Adv. Theor. Math. Phys.* **3** (1999) 209.

[29] J. C. Baez, J. D. Christensen, and G. Egan, "Asymptotics of 10j symbols," *Class. Quant. Grav.* **19** (2002) 6489.

[30] J. W. Barrett, R. J. Dowdall, W. J. Fairbairn, H. Gomes, and F. Hellmann, "Asymptotic analysis of the Engle-Pereira-Rovelli-Livine four-simplex amplitude," *J. Math. Phys.* **50** (2009) 112504, 0902.1170.

[31] J. W. Barrett, R. J. Dowdall, W. J. Fairbairn, H. Gomes, and F. Hellmann, "A summary of the asymptotic analysis for the EPRL amplitude," in *American Institute of Physics Conference Series*, J. Kowalski-Glikman, R. Durka, & M. Szczachor, ed., vol. 1196 of *American Institute of Physics Conference Series*, pp. 36–43. 2009. 0909.1882.

[32] J. W. Barrett, R. J. Dowdall, W. J. Fairbairn, H. Gomes, F. Hellmann, and R. Pereira, "Asymptotics of 4d spin foam models," *ArXiv e-prints* (2010) 1003.1886.

[33] J. W. Barrett, W. J. Fairbairn, and F. Hellmann, "Quantum gravity asymptotics from the $SU(2)$ 15j-symbol," *Int. J. Mod. Phys. A* **25** (2010) 2897–2916, 0912.4907.

[34] J. W. Barrett, R. J. Dowdall, W. J. Fairbairn, F. Hellmann, and R. Pereira, "Lorentzian spin foam amplitudes: graphical calculus and asymptotics," *Class. Quant. Grav.* **27** (2010) 165009, 0907.2440.

[35] J. Engle, E. Livine, R. Pereira, and C. Rovelli, "LQG vertex with finite Immirzi parameter," *Nuc. Phys. B* **799** (2008) 136–149, 0711.0146.

[36] L. Freidel and K. Krasnov, "A new spin foam model for 4D gravity," *Class. Quant. Grav.* **25** (2008), no. 12, 125018, 0708.1595.

[37] V. Aquilanti, H. M. Haggard, R. G. Littlejohn, and L. Yu, "Semiclassical analysis of Wigner 3j-symbol," *J. Phys. A* **40** (2007), no. 21, 5637.

[38] R. G. Littlejohn, "The Van Vleck formula, Maslov theory, and phase space geometry," *J. Stat. Phys.* **68** (1992) 7–50.

[39] G. Naber, *Topology, Geometry, and Gauge Fields: Foundations*. Springer, 1997.

[40] H. Koçak, F. Bisshopp, T. Banchoff, and D. Laidlaw, "Topology and mechanics with computer graphics: Linear Hamiltonian systems in four dimensions," *Adv. App. Math.* **7** (1986), no. 3, 282 – 308.

[41] M. V. Berry and K. E. Mount, "Semiclassical approximations in wave mechanics," *Rep. Prog. Phys.* **35** (1972) 315–397.

[42] I. C. Percival, "Semiclassical theory of bound states," *Adv. Chem. Phys.* **36** (1977) 1–61.

[43] S. Bates and A. Weinstein, *Lectures on the Geometry of Quantization*. Berkeley Mathematical Lecture Notes, 1997.

[44] J. Lansey, "Analysis of a single pendulum harmonograph." `http://jonathan.lansey.net/pastimes/pendulum/index.html`.

[45] R. Doll and G. Ingold, "Lissajous curves and semiclassical theory: The two-dimensional harmonic oscillator," *Am. J. Phys.* **75** (2007), no. 3, 208–215.

[46] H. M. Haggard, "Conservation, symmetry and Noether." Unpublished undergraduate thesis, 2002.

[47] N. W. Evans, "Superintegrability in classical mechanics," *Phys. Rev. A* **41** (May, 1990) 5666–5676.

[48] E. G. Kalnins, W. Miller, Jr., and G. S. Pogosyan, "Superintegrability and associated polynomial solutions: Euclidean space and the sphere in two dimensions," *J. Math. Phys.* **37** (1996) 6439–6467.

[49] G. Wentzel, "Eine verallgemeinerung der quantenbedingungen für die zwecke der wellenmechanik," *Zeit. Phys.* **38** (1926) 518–529.

[50] H. A. Kramers, "Wellenmechanik und halbzhlige Quantisierung," *Zeit. Phys.* **39** (1926) 828 – 840.

[51] L. Brillouin, "La mécanique ondulatoire de Schrödinger: une méthode générale de resolution par approximations successives," *Comp. Rend. Acad. Sci.* **183** (1926) 24–26.

[52] H. Jeffreys, "On certain approximate solutions of linear differential equations of the second order," *Proc. Lond. Math. Soc.* **23** (1924) 428–436.

[53] J. Boyd, "The Devil's Invention: Asymptotic, Superasymptotic and Hyperasymptotic Series," *Acta App. Math.* **56** (1999) 1–98.

[54] E. Schrödinger, "Quantisierung als Eigenwertproblem (II)," *Ann. Phys.* **79** (1926) 489–527.

[55] G. Ludwig, *Wave Mechanics*. Pergamon Press, 1968.

[56] V. I. Arnold, *Mathematical Methods of Classical Mechanics*. Springer, 1989.

[57] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry*. Springer, 1999.

[58] R. Abraham and J. E. Marsden, *Foundations of Mechanics*. Perseus, 1994.

[59] J. Ehlers and E. T. Newman, "The theory of caustics and wave front singularities with physical applications," *J. Math. Phys.* **41** (2000), no. 6, 3344–3378.

[60] A. Einstein, "Zum quantensatz von Sommerfeld und Epstein," *Dtsch. Phys. Ges.* **19** (1917) 82–92.

[61] L. Brillouin, "Remarques sur la mécanique ondulatoire," *J. Phys. Radium* **7** (1926) 353–368.

[62] J. B. Keller, "Corrected Bohr-Sommerfeld quantum conditions for non-separable systems," *Ann. Phys.* **4** (1958) 180–188.

[63] R. G. Littlejohn, "Semiclassical structure of trace formulas," *J. Math. Phys.* **31** (1990), no. 12, 2952–2977.

[64] D. J. Griffiths, *Introduction to Quantum Mechanics (2nd ed.)*. Prentice Hall, 2004.

[65] P. Ehrenfest, "On adiabatic changes of a system in connection with the quantum theory," *Proc. Amst. Acad.* **19** (1916) 576597.

[66] D. Byrd and D. Mitchell, "Adiabatic Bohr-Sommerfeld calculations for the hydrogenic Stark effect," *Phys. Rev. A* **70** (Dec, 2004) 065401–1–065401–4.

[67] M. Courtney, N. Spellmeyer, H. Jiao, and D. Kleppner, "Classical, semiclassical, and quantum dynamics in the lithium Stark system," *Phys. Rev. A* **51** (May, 1995) 3604–3620.

[68] V. P. Maslov and M. V. Fedoriuk, *Semi-Classical Approximations in Quantum Mechanics*. Dordrecht: D. Reidel, 1981.

[69] A. S. Mishchenko, V. E. Shatalov, and B. Y. Sternin, *Lagrangian Manifolds and the Maslov Operator*. Berlin: Springer Verlag, 1997.

[70] H. M. Haggard and R. G. Littlejohn, "Asymptotics of the Wigner $9j$ symbol," *Class. Quant. Grav.* **27** (2010) 135010, `0912.5384`.

[71] E. Bianchi and H. M. Haggard, "Discreteness of the volume of space from Bohr-Sommerfeld quantization," *ArXiv e-prints* (2011) 1102.5439.

[72] A. Weinstein, "The geometry of momentum," *ArXiv e-prints* (Aug., 2002) arXiv:math/0208108.

[73] J. S. Schwinger, "On angular momentum," *U.S. Atomic Energy Commision* (1952).

[74] V. Bargmann, "On the representations of the rotation group," *Rev. Mod. Phys.* **34** (1962) 829–845.

[75] H. V. McIntosh, "Symmetry and degeneracy," in *Group Theory and its Applications*, E. M. Loebl, ed. Academic Press, 1971.

[76] L. C. Biedenharn, J. D. Louck, and P. A. Carruthers, *Angular Momentum in Quantum Physics: Theory and Application*. Cambridge University Press, 2009.

[77] R. Penrose and W. Rindler, *Spinors and Spacetime (vol. 1)*. Cambridge University Press, 1984.

[78] S. A. Huggett and K. P. Todd, *An Introduction to Twistor Theory*. Cambridge University Press, 1994.

[79] A. Weinstein, "The local structure of Poisson manifolds," *J. Diff. Geom.* **18** (1983) 523–557.

[80] P. Olver, *Classical Invariant Theory*. Cambridge University Press, 1999.

[81] B. Sturmfels, *Algorithms in Invariant Theory*. SpringerWienNewYork, 2008.

[82] I. M. Gelfand, M. M. Kapranov, and A. Zelevinsky, *Discriminants, Resultants, and Multidimensional Determinants*. Birkhäuser, 2008.

[83] S. Major, *q-Quantum Gravity*. PhD thesis, The Pennsylvania State University, 1997.

[84] K. Reidemeister, "Elementare begründung der knotentheorie," *Abh. Math. Sem. Univ. Hamburg* **5** (1926) 24–32.

[85] C. M. Cramlet, "The derivation of algebraic invariants by tensor algebra," *Bull. Amer. Math. Society* **34** (1928) 334–342.

[86] D. Littlewood, "Invariant theory, tensors, and group characters," *Phil. Trans. Royal Soc.* **239** (1944) 305–365.

[87] G. B. Gurevich, *Foundations of the Theory of Algebraic Invariants*. P. Noordhorff Ltd., 1964.

[88] N. Jeevanjee, *An Introduction to Tensors for Physicists*. To be published by Birkhäuser, 2011.

[89] P. G. Appleby, B. R. Duffy, and R. W. Ogden, "On the classification of isotropic tensors," *Glasgow Math. J.* **29** (1987) 185–196.

[90] L. H. Kauffman, *Knots and Physics*. World Scientific, 2001.

[91] G. E. Stedman, *Diagram Techniques in Group Theory*. Cambridge University Press, 1990.

[92] H. Weyl, *The Classical Groups, (2nd ed. 1953)*. Princeton University Press, 1997.

[93] L. M. Blumenthal, *Theory and Applications of Distance Geometry*. Chelsea Publishing Co., 1970.

[94] G. M. Crippen and T. F. Havel, *Distance Geometry and Molecular Conformation*. Research Studies Press, 1988.

[95] J. Blinn, "Quartic discriminants and tensor invariants," *Comp. Graph. App., IEEE* **22** (2002), no. 2, 86–91.

[96] B. Sturmfels, *Solving Systems of Polynomial Equations*. American Mathematical Society, 2002.

[97] V. Dolotin and A. Moroz, *Introduction to Non-linear Algebra*. World Scientific, 2007.

[98] A. Hedeman and R. G. Littlejohn, "The 2j-symbol, Wigner rotation matrices, and their asymptotics." In preperation, 2011.

[99] A. R. Edmonds, *Angular Momentum in Quantum Mechanics*. Princeton University Press, 1996.

[100] E. El Baz and B. Castel, *Graphical Methods of Spin Algebras in Atomic, Nuclear and Particle Physics*. Dekker, 1972.

[101] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum*. World Scientific, 1981.

[102] P. E. S. Wormer and J. Paldus, "Angular momentum diagrams," *Adv. Quantum Chem.* **51** (2006) 59–123.

[103] P. Cvitanović, *Group Theory: Birdtracks, Lie's and Exceptional Groups*. Princeton University Press, 2008.

[104] L. H. Kauffman and S. Lins, *Temperley-Lieb Recoupling Theory and Invariants of 3-Manifolds*. Princeton University Press, 1994.

[105] J. P. Moussouris, *Quantum Models as Spacetime based on Recoupling Theory (unpublished)*. PhD thesis, Oxford, 1983.

[106] A. P. Yutsis, I. B. Levinson, and V. V. Vanagas, *The Theory of Angular Momentum*. Israel Program for Scientific Translations, 1962.

[107] D. M. Brink and G. R. Satchler, *Angular Momentum*. Oxford University Press, 1993.

[108] J. J. Sakurai, *Modern Quantum Mechanics*. Addison-Wesley, 1994.

[109] E. Bianchi, P. Doná, and S. Speziale, "Polyhedra in loop quantum gravity," *Phys. Rev. D* **83** (2011), no. 4, 044035, 1009.3402.

[110] M. Kapovich and J. J. Millson, "The symplectic geometry of polygons in Euclidean space," *J. Diff. Geom.* **44** (1996) 479–513.

[111] A. Chakrabarti, "On the coupling of 3 angular momenta," *Ann. Henri Poincaré* **1** (1964) 301–327.

[112] J. Lévy-Leblond and M. Lévy-Nahas, "Symmetrical coupling of three angular momenta," *J. Math. Phys.* **6** (1965) 1372–1380.

[113] R. de Pietri and C. Rovelli, "Geometry eigenvalues and the scalar product from recoupling theory in loop quantum gravity," *Phys. Rev. D* **54** (1996) 2664–2690, arXiv:gr-qc/9602023.

[114] T. Thiemann, "Closed formula for the matrix elements of the volume operator in canonical quantum gravity," *J. Math. Phys.* **39** (1998) 3347–3371, arXiv:gr-qc/9606091.

[115] G. Carbone, M. Carfora, and A. Marzuoli, "Quantum states of elementary three-geometry," *Classical and Quantum Gravity* **19** (2002) 3761–3774, arXiv:gr-qc/0112043.

[116] K. A. Meissner, "Eigenvalues of the volume operator in loop quantum gravity," *Class. Quant. Grav.* **23** (2006) 617–625, arXiv:gr-qc/0509049.

[117] J. Brunnemann and T. Thiemann, "Simplification of the spectral analysis of the volume operator in loop quantum gravity," *Classical and Quantum Gravity* **23** (2006) 1289–1346, arXiv:gr-qc/0405060.

[118] J. Brunnemann and D. Rideout, "Properties of the volume operator in loop quantum gravity: I. Results," *Classical and Quantum Gravity* **25** (2008), no. 6, 065001, 0706.0469.

[119] J. Brunnemann and D. Rideout, "Properties of the volume operator in loop quantum gravity: II. Detailed presentation," *Classical and Quantum Gravity* **25** (2008), no. 6, 065002, 0706.0382.

[120] H. Minkowski, "Allgemeine Lehrsätze über die konvexe polyeder," *Nachr. Ges. Wiss.* **Göttingen** (1897) 198–219.

[121] C. Rovelli and L. Smolin, "Discreteness of area and volume in quantum gravity," *Nuc. Phys. B* **442** (1995) 593–619, arXiv:gr-qc/9411005.

[122] A. Ashtekar and J. Lewandowski, "Quantum Theory of Geometry II: Volume operators," *Adv. Theor. Math. Phys.* **1** (1998) 388–429, arXiv:gr-qc/9711031.

[123] J. Brunnemann and D. Rideout, "Oriented matroids—combinatorial structures underlying loop quantum gravity," *Classical and Quantum Gravity* **27** (2010) 205008, 1003.2348.

[124] C. Flori and T. Thiemann, "Semiclassical analysis of the Loop Quantum Gravity volume operator: I. flux coherent states," *ArXiv e-prints* (2008) 0812.1537.

[125] A. Barbieri, "Quantum tetrahedra and simplicial spin networks," *Nuclear Physics B* **518** (1998) 714–728, arXiv:gr-qc/9707010.

[126] S. A. Major, "Operators for quantized directions," *Classical and Quantum Gravity* **16** (1999) 3859–3877, arXiv:gr-qc/9905019.

[127] N. Grot and C. Rovelli, "Moduli-space structure of knots with intersections," *J. Math. Phys.* **37** (1996) 3014–3021, gr-qc/9604010.

[128] J. B. Lasserre, "An analytical expression and an algorithm for the volume of a convex polyhedron in R^n," *J. Optim. Theor. Appl.* **39** (1983) 363–377.

[129] S. A. Major, "Shape in an atom of space: exploring quantum geometry phenomenology," *Class. Quant. Grav.* **27** (2010), no. 22, 225012–+, 1005.5460.

[130] A. Ashtekar and J. Lewandowski, "Background independent quantum gravity: a status report," *Classical and Quantum Gravity* **21** (2004) 53, arXiv:gr-qc/0404018.

[131] C. Rovelli, "Loop quantum gravity: the first twenty five years," *ArXiv e-prints* (2010) 1012.4707.

[132] M. C. Gutzwiller, *Chaos in Classical and Quantum Mechanics*. Springer, 1990.

[133] M. V. Berry, "Waves and Thom's theorem," *Adv. Phys.* **25** (1976) 1–26.

[134] L. C. Biedenharn and J. D. Louck, *The Racah-Wigner Algebra in Quantum Theory*. Cambridge University Press, 2009.

[135] R. W. Anderson, V. Aquilanti, and C. da S. Ferreira, "Exact computation and large angular momentum asymptotics of $3nj$-symbols: Semiclassical disentangling of spin networks," *J. Chem. Phys.* **129** (2008) 161101.

[136] R. W. Anderson, V. Aquilanti, and A. Marzuoli, "$3nj$ morphogenesis and semiclassical disentangling," *J. Phys. Chem.* **113** (2009) 15106.

[137] A. Marzuoli and M. Rasetti, "Computing spin networks," *Ann. Phys.* **318** (2005), no. 2, 345–407.

[138] E. P. Wigner, *Group Theory*. Academic, 1959.

[139] K. Schulten and R. G. Gordon, "Exact recursive evaluation of $3j$- and $6j$-coefficients for quantum-mechanical coupling of angular momenta," *J. Math. Phys.* **16** (1975), no. 10, 1961–1970.

[140] K. Schulten and R. G. Gordon, "Semiclassical approximations to $3j$- and $6j$-coefficients for quantum-mechanical coupling of angular momenta," *J. Math. Phys.* **16** (1975), no. 10, 1971–1988.

[141] H. Trinkaus and F. Drepper, "On the analysis of diffraction catastrophes," *J. Math. Phys. A* **10** (1977), no. 1, L11.

[142] G. T. Gilbert, "Positive definite matrices and Sylvester's criterion," *Amer. Math. Monthly* **98** (1991), no. 1, 44–46.

[143] J. Sommese and C. W. I. Wampler, *The Numerical Solution of Systems of Polynomials arising in Engineering and Science*. World Scientific, 2005.

[144] M. Kapovich, J. J. Millson, and T. Treloar, "The symplectic geometry of polygons in hyperbolic 3-space," *Asian J. Math.* **4** (2000) 123–164.

[145] R. G. Littlejohn and L. Yu, "Semiclassical Analysis of the Wigner \$9J\$-Symbol with small and large angular momenta," *ArXiv e-prints* (2011) 1104.1499.

[146] L. Yu, "Semiclassical analysis of the Wigner \$12J\$-symbol with one small angular momentum: Part I," *ArXiv e-prints* (2011) 1104.3275.

[147] L. Yu, "Asymptotic limits of the Wigner \$15J\$-symbol with small quantum numbers," *ArXiv e-prints* (Apr., 2011) 1104.3641.

[148] J. W. Barrett, "First order Regge calculus," *Class. Quant. Grav.* **11** (1994) 2723–2730, arXiv:hep-th/9404124.

[149] B. Bahr and B. Dittrich, "Regge calculus from a new angle," *New J. Phys.* **12** (2010), no. 3, 033010, 0907.4325.

[150] R. W. Sharpe, *Differential Geometry: Cartan's Generalization of Klein's Erlangen Program*. Springer, 1997.

[151] J. V. José and E. J. Saletan, *Classical Dynamics: A Contemporary Approach*. Cambridge University Press, 1998.

[152] T. Padmanabhan, *Gravitation: Foundations and Frontiers*. Cambridge University Press, 2010.

[153] R. V. Levine, *A Geography Of Time: On Tempo, Culture, And The Pace Of Life*. Basic Books, 1997.

[154] L. M. A. Bettencourt, J. Lobo, D. Helbing, C. Kühnert, and G. B. West, "Growth, innovation, scaling, and the pace of life in cities," *PNAS* **104** (2007), no. 17, 7301–7306.

Appendix A

Root Series for the Volume Operator

The definitions that this appendix relies on are given sections 4.4 and 4.5 of Chapter 4. For the case where the $\bar{r}_i$ are distinct we have,

$$r_i = \bar{r}_i + \lambda_i Q^2 + \beta_i \lambda_i^2 Q^4 + \cdots , \tag{A.1}$$

where

$$\lambda_i \equiv \frac{18^2 \bar{r}_i}{\prod_{j\neq i}(\bar{r}_i - \bar{r}_j)} \qquad \text{and} \qquad \beta_i \equiv \left(\frac{1}{\bar{r}_i} - \sum_{j\neq i} \frac{1}{\bar{r}_i - \bar{r}_j} \right). \tag{A.2}$$

When $\bar{r}_1$ and $\bar{r}_2$ coincide, say at $\bar{r}$, the above series are singular and are instead replaced by,

$$r_i = \bar{r} \pm \mu Q + \nu Q^2 \pm \rho Q^3 + \sigma Q^4 \pm \cdots \qquad (i = 1, 2), \tag{A.3}$$

the lower signs for r_1 and the upper for r_2 and with,

$$\mu = \frac{18\sqrt{\bar{r}}}{\sqrt{(\bar{r}_3 - \bar{r})(\bar{r}_4 - \bar{r})}}, \qquad \nu = \frac{(18)^2(\bar{r}_4\bar{r}_3 - \bar{r}^2)}{2(\bar{r}_3 - \bar{r})^2(\bar{r}_4 - \bar{r})^2}, \tag{A.4}$$

$$\rho = \frac{(18)^3(\bar{r}_3^2\bar{r}_4^2 + 4\bar{r}\bar{r}_3\bar{r}_4(\bar{r}_3 + \bar{r}_4) - 14\bar{r}^2\bar{r}_3\bar{r}_4 + 5\bar{r}^4)}{8\sqrt{\bar{r}}(\bar{r}_3 - \bar{r})^{7/2}(\bar{r}_4 - \bar{r})^{7/2}}, \tag{A.5}$$

$$\sigma = \frac{(18)^4(\bar{r}_3^2\bar{r}_4^2(\bar{r}_3 + \bar{r}_4) + \bar{r}\bar{r}_3\bar{r}_4(\bar{r}_3 - \bar{r}_4)^2 - 5\bar{r}^2\bar{r}_3\bar{r}_4(\bar{r}_3 + \bar{r}_4) + 10\bar{r}^3\bar{r}_3\bar{r}_4 - 2\bar{r}^5)}{2(\bar{r}_3 - \bar{r})^5(\bar{r}_4 - \bar{r})^5}. \tag{A.6}$$

If $\bar{r}_3$ and $\bar{r}_4$ coincide, say at $\tilde{r}$, the above formula holds with the replacements $\bar{r}_3 \to \bar{r}_1$, $\bar{r}_4 \to \bar{r}_2$ and $\bar{r} \to \tilde{r}$. There is one more set of singular cases: when $\bar{r}_1$ is zero then r_1 is zero for all Q and the series expansions of the other roots changes. For $\bar{r}_2$, $\bar{r}_3$ and

$\bar{r}_4$ distinct these are,

$$r_j = \bar{r}_j + \frac{(18)^2}{\prod_{k\neq j}(\bar{r}_j - \bar{r}_k)}Q^2 + \frac{(18)^4 \sum_{k\neq j}(\bar{r}_k - \bar{r}_j)}{\prod_{k\neq j}(\bar{r}_j - \bar{r}_k)^3}Q^4 + \cdots \qquad (j = 2, 3, 4). \quad \text{(A.7)}$$

This series is well behaved when $\bar{r}_2$ is zero, *i.e.* when $\bar{r}_1 = \bar{r}_2$, but singular when $\bar{r}_3 = \bar{r}_4 = \tilde{r}$ and so we have one final case when $r_1 = 0$ and the larger roots coalesce:

$$r_2 = \bar{r}_2 + \frac{(18)^2}{(\tilde{r} - \bar{r}_2)^2}Q^2 + \frac{2(18)^4}{(\tilde{r} - \bar{r}_2)^5}Q^4 + \cdots, \quad \text{(A.8)}$$

$$r_k = \bar{r}_k \pm \frac{18}{\sqrt{\tilde{r} - \bar{r}_2}}Q - \frac{(18)^2}{2(\tilde{r} - \bar{r}_2)^2}Q^2 \pm \frac{5(18)^3}{8(\tilde{r} - \bar{r}_2)^{7/2}}Q^3 - \frac{(18)^4}{(\tilde{r} - \bar{r}_2)^5}Q^4 \cdots \qquad (k = 3, 4), \quad \text{(A.9)}$$

with the lower signs for r_3 and the upper signs for r_4.

Appendix B

Adjacency and the Minkowski reconstruction

In this appendix we present an exhaustive list of the closure relations for a polyhedron with five faces. These closure relations are also algebraically related to the parameters $\alpha = \alpha_1, \beta = \beta_1$ and $\gamma = \gamma_1$ of the first case:

1. $\alpha_1 \boldsymbol{A}_1 + \beta_1 \boldsymbol{A}_2 + \gamma_1 \boldsymbol{A}_3 + \boldsymbol{A}_4 = 0$

$$\alpha \equiv \alpha_1 = -\frac{W_{234}}{W_{123}} \qquad \beta \equiv \beta_1 = \frac{W_{134}}{W_{123}} \qquad \gamma \equiv \gamma_1 = -\frac{W_{124}}{W_{123}}. \tag{B.1}$$

2. $\alpha_2 \boldsymbol{A}_1 + \beta_2 \boldsymbol{A}_2 + \boldsymbol{A}_3 + \gamma_2 \boldsymbol{A}_4 = 0$

$$\alpha_2 = \frac{W_{234}}{W_{124}} = \frac{\alpha}{\gamma} \qquad \beta_2 = -\frac{W_{134}}{W_{124}} = \frac{\beta}{\gamma} \qquad \gamma_2 = -\frac{W_{123}}{W_{124}} = \frac{1}{\gamma}. \tag{B.2}$$

3. $\alpha_3 \boldsymbol{A}_1 + \boldsymbol{A}_2 + \beta_3 \boldsymbol{A}_3 + \gamma_3 \boldsymbol{A}_4 = 0$

$$\alpha_3 = -\frac{W_{234}}{W_{134}} = \frac{\alpha}{\beta} \qquad \beta_3 = -\frac{W_{124}}{W_{134}} = \frac{\gamma}{\beta} \qquad \gamma_3 = \frac{W_{123}}{W_{134}} = \frac{1}{\beta}. \tag{B.3}$$

4. $\boldsymbol{A}_1 + \alpha_4 \boldsymbol{A}_2 + \beta_4 \boldsymbol{A}_3 + \gamma_4 \boldsymbol{A}_4 = 0$

$$\alpha_4 = -\frac{W_{134}}{W_{234}} = \frac{\beta}{\alpha} \qquad \beta_4 = \frac{W_{124}}{W_{234}} = \frac{\gamma}{\alpha} \qquad \gamma_4 = -\frac{W_{123}}{W_{234}} = \frac{1}{\alpha}. \tag{B.4}$$

5. $\alpha_5 \boldsymbol{A}_1 + \beta_5 \boldsymbol{A}_2 + \gamma_5 \boldsymbol{A}_3 + \boldsymbol{A}_5 = 0$

$$\alpha_5 = -\frac{W_{235}}{W_{123}} = 1 - \alpha \qquad \beta_5 = \frac{W_{135}}{W_{123}} = 1 - \beta \qquad \gamma_5 = -\frac{W_{125}}{W_{123}} = 1 - \gamma. \tag{B.5}$$

6. $\alpha_6 \boldsymbol{A}_1 + \beta_6 \boldsymbol{A}_2 + \boldsymbol{A}_3 + \gamma_6 \boldsymbol{A}_5 = 0$

$$\alpha_6 = \frac{W_{235}}{W_{125}} = \frac{1-\alpha}{1-\gamma} \qquad \beta_6 = -\frac{W_{135}}{W_{125}} = \frac{1-\beta}{1-\gamma} \qquad \gamma_6 = -\frac{W_{123}}{W_{125}} = \frac{1}{1-\gamma}. \tag{B.6}$$

7. $\alpha_7 \boldsymbol{A}_1 + \boldsymbol{A}_2 + \beta_7 \boldsymbol{A}_3 + \gamma_7 \boldsymbol{A}_5 = 0$

$$\alpha_7 = -\frac{W_{235}}{W_{135}} = \frac{1-\alpha}{1-\beta} \qquad \beta_7 = -\frac{W_{125}}{W_{135}} = \frac{1-\gamma}{1-\beta} \qquad \gamma_7 = \frac{W_{123}}{W_{135}} = \frac{1}{1-\beta}. \tag{B.7}$$

8. $\boldsymbol{A}_1 + \alpha_8 \boldsymbol{A}_2 + \beta_8 \boldsymbol{A}_3 + \gamma_8 \boldsymbol{A}_5 = 0$

$$\alpha_8 = -\frac{W_{135}}{W_{235}} = \frac{1-\beta}{1-\alpha} \qquad \beta_8 = \frac{W_{125}}{W_{235}} = \frac{1-\gamma}{1-\alpha} \qquad \gamma_8 = -\frac{W_{123}}{W_{235}} = \frac{1}{1-\alpha}. \tag{B.8}$$

9. $\alpha_9 \boldsymbol{A}_1 + \beta_9 \boldsymbol{A}_2 + \gamma_9 \boldsymbol{A}_4 + \boldsymbol{A}_5 = 0$

$$\alpha_9 = -\frac{W_{245}}{W_{124}} = 1 - \frac{\alpha}{\gamma} \qquad \beta_9 = \frac{W_{145}}{W_{124}} = 1 - \frac{\beta}{\gamma} \qquad \gamma_9 = -\frac{W_{125}}{W_{124}} = 1 - \frac{1}{\gamma}. \tag{B.9}$$

10. $\alpha_{10} \boldsymbol{A}_1 + \beta_{10} \boldsymbol{A}_2 + \boldsymbol{A}_4 + \gamma_{10} \boldsymbol{A}_5 = 0$

$$\alpha_{10} = \frac{W_{245}}{W_{125}} = \frac{\gamma-\alpha}{\gamma-1} \qquad \beta_{10} = -\frac{W_{145}}{W_{125}} = \frac{\gamma-\beta}{\gamma-1} \qquad \gamma_{10} = -\frac{W_{124}}{W_{125}} = \frac{\gamma}{\gamma-1}. \tag{B.10}$$

11. $\alpha_{11} \boldsymbol{A}_1 + \boldsymbol{A}_2 + \beta_{11} \boldsymbol{A}_4 + \gamma_{11} \boldsymbol{A}_5 = 0$

$$\alpha_{11} = -\frac{W_{245}}{W_{145}} = \frac{\gamma-\alpha}{\gamma-\beta} \qquad \beta_{11} = -\frac{W_{125}}{W_{145}} = \frac{\gamma-1}{\gamma-\beta} \qquad \gamma_{11} = \frac{W_{124}}{W_{145}} = \frac{\gamma}{\gamma-\beta}. \tag{B.11}$$

12. $\boldsymbol{A}_1 + \alpha_{12} \boldsymbol{A}_2 + \beta_{12} \boldsymbol{A}_4 + \gamma_{12} \boldsymbol{A}_5 = 0$

$$\alpha_{12} = -\frac{W_{145}}{W_{245}} = \frac{\gamma-\beta}{\gamma-\alpha} \qquad \beta_{12} = \frac{W_{125}}{W_{245}} = \frac{\gamma-1}{\gamma-\alpha} \qquad \gamma_{12} = -\frac{W_{124}}{W_{245}} = \frac{\gamma}{\gamma-\alpha}. \tag{B.12}$$

13. $\alpha_{13} \boldsymbol{A}_1 + \beta_{13} \boldsymbol{A}_3 + \gamma_{13} \boldsymbol{A}_4 + \boldsymbol{A}_5 = 0$

$$\alpha_{13} = -\frac{W_{345}}{W_{134}} = 1 - \frac{\alpha}{\beta} \qquad \beta_{13} = \frac{W_{145}}{W_{134}} = 1 - \frac{\gamma}{\beta} \qquad \gamma_{13} = -\frac{W_{135}}{W_{134}} = 1 - \frac{1}{\beta}. \tag{B.13}$$

14. $\alpha_{14}\boldsymbol{A}_1 + \beta_{14}\boldsymbol{A}_3 + \boldsymbol{A}_4 + \gamma_{14}\boldsymbol{A}_5 = 0$

$$\alpha_{14} = -\frac{W_{345}}{W_{135}} = \frac{\beta-\alpha}{\beta-1} \qquad \beta_{14} = -\frac{W_{145}}{W_{135}} = \frac{\beta-\gamma}{\beta-1} \qquad \gamma_{14} = -\frac{W_{134}}{W_{135}} = \frac{\beta}{\beta-1}. \tag{B.14}$$

15. $\alpha_{15}\boldsymbol{A}_1 + \boldsymbol{A}_3 + \beta_{15}\boldsymbol{A}_4 + \gamma_{15}\boldsymbol{A}_5 = 0$

$$\alpha_{15} = -\frac{W_{345}}{W_{145}} = \frac{\beta-\alpha}{\beta-\gamma} \qquad \beta_{15} = -\frac{W_{135}}{W_{145}} = \frac{\beta-1}{\beta-\gamma} \qquad \gamma_{15} = -\frac{W_{134}}{W_{145}} = \frac{\beta}{\beta-\gamma}. \tag{B.15}$$

16. $\boldsymbol{A}_1 + \alpha_{16}\boldsymbol{A}_3 + \beta_{16}\boldsymbol{A}_4 + \gamma_{16}\boldsymbol{A}_5 = 0$

$$\alpha_{16} = -\frac{W_{145}}{W_{345}} = \frac{\beta-\gamma}{\beta-\alpha} \qquad \beta_{16} = \frac{W_{135}}{W_{345}} = \frac{\beta-1}{\beta-\alpha} \qquad \gamma_{16} = -\frac{W_{134}}{W_{345}} = \frac{\beta}{\beta-\alpha}. \tag{B.16}$$

17. $\alpha_{17}\boldsymbol{A}_2 + \beta_{17}\boldsymbol{A}_3 + \gamma_{17}\boldsymbol{A}_4 + \boldsymbol{A}_5 = 0$

$$\alpha_{17} = -\frac{W_{345}}{W_{234}} = 1 - \frac{\beta}{\alpha} \qquad \beta_{17} = \frac{W_{245}}{W_{234}} = 1 - \frac{\gamma}{\alpha} \qquad \gamma_{17} = -\frac{W_{235}}{W_{234}} = 1 - \frac{1}{\alpha}. \tag{B.17}$$

18. $\alpha_{18}\boldsymbol{A}_2 + \beta_{18}\boldsymbol{A}_3 + \boldsymbol{A}_4 + \gamma_{18}\boldsymbol{A}_5 = 0$

$$\alpha_{18} = \frac{W_{345}}{W_{235}} = \frac{\alpha-\beta}{\alpha-1} \qquad \beta_{18} = -\frac{W_{245}}{W_{235}} = \frac{\alpha-\gamma}{\alpha-1} \qquad \gamma_{18} = -\frac{W_{234}}{W_{235}} = \frac{\alpha}{\alpha-1}. \tag{B.18}$$

19. $\alpha_{19}\boldsymbol{A}_2 + \boldsymbol{A}_3 + \beta_{19}\boldsymbol{A}_4 + \gamma_{19}\boldsymbol{A}_5 = 0$

$$\alpha_{19} = -\frac{W_{345}}{W_{245}} = \frac{\alpha-\beta}{\alpha-\gamma} \qquad \beta_{19} = -\frac{W_{235}}{W_{245}} = \frac{\alpha-1}{\alpha-\gamma} \qquad \gamma_{19} = -\frac{W_{234}}{W_{245}} = \frac{\alpha}{\alpha-\gamma}. \tag{B.19}$$

20. $\boldsymbol{A}_2 + \alpha_{20}\boldsymbol{A}_3 + \beta_{20}\boldsymbol{A}_4 + \gamma_{20}\boldsymbol{A}_5 = 0$

$$\alpha_{20} = -\frac{W_{245}}{W_{345}} = \frac{\alpha-\gamma}{\alpha-\beta} \qquad \beta_{20} = -\frac{W_{235}}{W_{345}} = \frac{\alpha-1}{\alpha-\beta} \qquad \gamma_{20} = -\frac{W_{234}}{W_{345}} = \frac{\alpha}{\alpha-\beta}. \tag{B.20}$$