Geometric Control For Analysing the Quantum Speed Limit and the Physical Limitations of Computers

Benjamin James Russell

Doctor of Philosophy

University Of York
Computer Science

February 2015

Abstract

This thesis studies the role of Finsler geometry in quantum time optimal control of systems with constrained control field power and other constraints. The systems considered are all finite dimensional systems with pure states. A Finsler metric is constructed such that its geodesics are the time optimal trajectories for the quantum time evolution operator on the special unitary group. This metric is shown to be right invariant. The geodesic equation, in the form of an Euler-Poincaré equation is found. It is also shown that the geodesic lengths of this same metric equal the optimal times for implementing any desired quantum gate. In a special case, where all are control fields are equally constrained, the desired geodesics are found in closed form. The results obtained are discussed in the general context of natural computation.

Contents

List of Figures

Acknowledgements

Firstly, I would like to thank my Supervisor Prof Susan Stepney for her excellent research guidance and encouragement and many interesting and helpful discussions not limited to the work in the thesis. Secondly, I would like to thanks Dr. Eli Hawkins and Dr. Ian McIntosh of the University of York Maths Department for their time in many helpful discussions about geometry. Thirdly, I would like to thank the reviewers of my journal papers for their extremely through comments and corrections. Finally I would like to thank my partner Ioulia Anastasiou for putting up with me complaining about writing this thesis and so patiently supporting me, and my parents and my brothers you also supported me during my Ph.D.

Author's Declaration

The work presented in this thesis is all my own except where explicitly indicated and cited. The work present in chapters 6, 7, 8 and 9 has been published in International Journal of the Foundations of Computer Science [1], Physical Review A [2] and Journal of Physics A [3]. This work has not previously been presented for an award at this, or any other, University.

Chapter 1

Introduction

1.1 Contribution

The central contribution of this thesis is to explicitly construct a right invariant Randers metric on the special unitary group (of arbitrary dimension):

$$F_{\hat{U}}(\hat{A}\hat{U}) = \sqrt{\frac{h(\hat{A},\hat{A})}{\lambda} + \frac{h(\hat{A},i\hat{H}_0)^2}{\lambda^2}} + \frac{h(\hat{A},i\hat{H}_0)}{\lambda} = F(\hat{A}) \tag{1.1}$$

for which the time optimal trajectories (of the time evolution operator $\hat{U}_t$) for implementing a quantum gate, using a limited amount of a resource to implement the control fields, are the geodesics. This result applies to all systems with Hamiltonian:

$$\hat{H}_t = \hat{H}_0 + \hat{H}_c(t) \tag{1.2}$$

with $\hat{H}_0$ time independent and, it a partucular sence, smaller than $\hat{H}_0$.

Furthermore it is shown how sub-Randers metrics are in one-to-one correspondence with a broad range of quantum control problems encountered in practice. These results are independent of the number of qubits to which they apply.

The geodesics in the cases that the limited resourse constraint is $\kappa \operatorname{Tr}\left(\hat{H}_c(t)^2\right) = 1$ are found explicitly and so are the control Hamiltonians driving the system along such geodesics. The geodesics are shown to each be of the form of the product of two one parameter subgroups:

$$\hat{U}_t = \exp\left(-it\hat{H}_0\right)\exp\left(it\hat{D}\right) \tag{1.3}$$

for some $i\hat{D} \in \mathfrak{su}(n)$. A formula is also found the choice of $\hat{D}$ which causes the system to impliment the gate $\hat{O}$:

$$\hat{D} = \frac{1}{T}\log\left(\exp\left(iT\hat{H}_0\right)\hat{O}\right) \tag{1.4}$$

where T is the time required to do so. It is shown that T satisfies a complex implicit equation. This equation is sovled numerically for some simple examples. The optimal control Hamiltonian driving the system along a geodesic is shown to be:

$$\hat{H}_c(t) = \frac{i}{T}\exp\left(-it\hat{H}_0\right)\log\left(\exp\left(iT\hat{H}_0\right)\hat{O}\right)\exp\left(it\hat{H}_0\right) \tag{1.5}$$

The technique of Euler-Poincare reduction is applied to the metric above for arbitrary constraints. This results on a first order differential equation, albeit a complicated one, for the optimal Hamiltonian driving the system along a geodesic.

This geometric method is applied to several few qubit systems and they are analysed numbericaly. It is also shown how this method has an advantage over the method presented in [4, 5] as it can prove optimal times for arbitrary trajectories, not just geodesics.

State control problems are also discussed alongside the problem of time optimally implementing quantum gates. It is argued that these two problems possess a strong correspondence and that the former can always be cast as an example of the latter. It is further argued that this method has the potential for broader application in physics based models of computation, the limitations of this are also discussed.

1.2 Project Statement

The task of understanding the physical basis of computation was really first posed by Richard Feynman in his seminal lectures on computation [6] and 'There's Plenty of Room at the Bottom' [7]. This lecture posed several questions about the capacity and limitations of molecular scale machines. Following this, the goal of this thesis is broadly 'to better understand the physical limitations of computation'. However, this is an extremely broad question. Physics based models of computation are numerous [8, 9, 10, 11] and various in nature, thus a representative example model must be chosen before proceeding if we are to avoid attempting a task which is impractically general.

Quantum mechanics is a vitally important part of modern physics and any physically meaningful theory of computation must take it into account [12]. Quantum mechanics also has a clear, well establish and rigorous mathematical basis which is highly amenable to applications of geometric control theory, a well establish and powerful tool for determining optimal times in controlled systems of the type typically applied in quantum computing [13]. Geometric control theory is, unfortunately, not well adapted for easily understanding optimal control of relativistic systems or even infinite dimensional ones. Typically, the assumptions of geometric control theory include that:

- time is an independent parameter
- the state space of the system under control is a finite dimensional smooth manifold

There are alternative ways to study the control of systems not meeting these premises. There have been some investigation into geometric control theory for infinite dimensional controlled systems [14, 15]. These are however, highly mathematically involved. There have also been a few investigations into the time optimal control of relativistic systems [16, 17, 18] but these ideas are far from as developed as the non-relativistic cases. As such there is not such a standardised methodology for such problems. Before one can assess the limitations of relativistic computing devices, first a canonical theory of computation embodied in such systems is required. Investigations into relativistic quantum computation have been recently initiated [19, 20] along with many other papers from the same group and others based in Nottingham, Vienna, Warsaw and others. However, thus far the most currently practically plausible and testable theory of physics based computation is non-relativistic quantum computation.

In this thesis the specific issue of finding the optimal physical times for implementing quantum gates in finite dimensional, non relativistic quantum systems using bounded resources is addressed. The purpose of this choice is multi-faceted. One is that it is an interesting an important theoretical and practical question for quantum computing. However, more is sought. A geometric methodology for assessing the physical limitations of computation in terms of time optimal control is the true target. It is the hope that the novel methods described in this thesis could form part of the basis for such work and that, mutatis mutandis, they could be applied to physics based theories of computation along side non-relativistic quantum computation. This possibility is discussed in 12.

1.3 Assumed Knowledge

A good deal of mathematical background is assumed as is a large amount of computer science background. The assumed knowledge includes much of basic linear algebra, group theory, Lie theory, and some elementary differential equations. The basics of differential geometry, Riemannian geometry, metric spaces and topology are also assumed.

All the specialist concepts are developed in the text. The relevant aspects of Finsler geometry and Zermelo Navigation are clearly outlined in a self contained manner. Some of the basics of theoretical computer science are also assumed knowledge. Familiarity with the introductory concepts of algorithms and time complexity is essential.

Part I

Literature Review and Critique

Chapter 2

Mathematics of Zermelo Navigation

In this chapter, Zermelo's navigation problem is specified with a little historical detail. Shen's inspired solution to the problem using Finsler geometry is also described. Some of the required mathematics of Finsler geometry and, more specifically, Randers geometry is also presented. Riemannian geometry is not discussed in detail as there are so many standard references which give a far more through presentation than would be possible in this thesis.

The purpose of reviewing this material is to set up the application of Shen's theorem to quantum time optimal quantum control.

2.1 Metric Structures

In the following we use the notation T_pM to refer to the tangent space at the point p on a manifold M, and TM to refer to the entire tangent bundle of a manifold, that is, all tangent spaces considered together. For example $T_{\hat{U}}SU(n)$ is the tangent space to $SU(n)$ at the point $\hat{U} \in SU(n)$. We use the notation $\Gamma(TM)$ to refer to the set of all smooth sections of the tangent bundle TM, that is, effectively to say, all smooth vector fields on M. For good references using this notation for both general manifolds and Lie groups, see [21, 22].

2.1.1 Minkowski Norms

A Minkowski norm $F : V \to \mathbb{R}$ on a finite dimensional, real vector space $V \cong \mathbb{R}^n$ is a slight generalisation of a norm is the usual sense. A Minkowski norm $F : V \to \mathbb{R}$ must satisfy ([23], ch 1):

- $F(v) > 0 \;\; \forall v \in V \setminus \{0\}$ and $(F(v) = 0 \;\; \Leftrightarrow v = 0)$. Positivity.
- $F(\lambda v) = \lambda F(v) \;\; \forall \lambda \in \mathbb{R}^+$. Positivity Homogeneity, N.B. this is different to a norm.
- $F(u + v) \leq F(u) + F(v)$. Triangle inequality.

Note that $F(v) \neq F(-v)$ in general. This property is known as *reversibility* of a Minkowski norm. One can show that a reversible Minkowski norm is always a norm in the regular sense.

2.1.2 Finsler Metrics

A Finsler metric is a generalisation of a Riemannian metric on a manifold. A Finsler metric on a manifold M can be defined as a function $F_p : T_pM \to \mathbb{R}$ which is smooth on the slit tangent bundle $TM \setminus \{0\}$ and is a Minkowski norm on each tangent space T_pM.

Riemannian Examples

All Riemannian metrics g on a manifold M are examples of Finsler metrics. This can be seen by setting:

$$F_p(v) = \sqrt{g_p\,(v, v)} \tag{2.1}$$

for all $v \in T_pM$.

Non-Riemannian Example

An example of a Finsler metric on the plane $F_p : T_p\mathbb{R}^2 \to \mathbb{R}$ is the l-q norm on each tangent space, with q given by a not necessarily constant smooth function of p.

$$F_p(v) = \sqrt[q]{(v^0)^q + (v^1)^q} \tag{2.2}$$

2.1.3 Randers Metrics

Prominent examples of Finlser metrics are the Randers metrics. These were originally introduced by Randers in the context of the motion of particles in an electromagnetic field in general relativity [23, 24], but since they have found many applications in diverse fields [25, 26]. It was in fact [27] who first realised that the metric inroduced by Randers was in fact a Finsler metric, this article also was the first to name them Randers metrics.

To define a Randers metric, first one needs to define a Randers norm.

Definition 2.1.1. A *Randers norm* on a vector space $V \cong \mathbb{R}^N$ is a Minkowski norm on $F : V \to \mathbb{R}$ which takes the form:

$$F(v) = \sqrt{\alpha(v, v)} + \beta(v) \tag{2.3}$$

where α is an inner product and β is a one form. To ensure the positivity of the metric F it is also required that β be 'small' according to α in the sense that $\alpha(\beta^\sharp, \beta^\sharp) < 1$.

The $\sharp$ here is intended to indicate the musical isomorphism $V^* \to V$ induced by α. For the definition of $\flat, \sharp$ see any good textbook on Riemannian geometry which discusses the 'musical isomorphism' including ([28] §2.66). In coordinates this can also be expressed as $\alpha_{ij}\beta^i\beta^j < 1$.

Definition 2.1.2 (Randers Metric)**.** A *Randers metric* on a manifold M is a Finsler metric which is a Randers norm on each tangent space T_pM.

2.2 Zermelo's Navigation Problem

Zermelo's navigation problem is the problem of finding time optimal trajectories on a manifold M under certain restrictions. The first restriction is that the navigator has some constraint on the speed they can travel at 'under their own steam' alone. The second restriction is that there is a vector field W (for 'wind') on M pushing the navigator around. The original motivation for the problem was that of time optimally navigating a ship on a windy sea or a zeppelin in the wind.

2.2.1 Formal Statement of Problem

An exact statement of the general problem can be found in [29]. For a historical reference see [30].

To specify a navigation problem one needs:

- A smooth manifold M
- A Finsler metric G on M
- A smooth 'wind' or 'drift' vector field $W \in \Gamma(TM)$ such that $G_p(W_p) < 1 \ \forall p \in M$

Definition 2.2.1. Together (G, W) are known as the *Navigation data* for a Zermelo navigation problem on M.

The interpretation of G is that the navigator is constrained to have speed 1 according to G. That is to say, neglecting the effects of the 'wind' W on the navigators speed, all admissible trajectories $\gamma(t)$ satisfy $G\left(\frac{d\gamma}{dt}\right) = 1 \ \forall t$. This represents the limitation of the navigator's 'engine power' in the case of a ship on the sea type scenario.

The wind is interpreted as 'blowing' the navigator around as they attempt to navigate. This deforms the optimal trajectories. In the absence of wind it is clear that the optimal trajectories are be the geodesics of G with desired end-points. The wind affects the navigator in the sense that the allowed set of tangent vectors to a trajectory passing through the point $p \in M$ has the form $\frac{d\gamma}{dt} = W_{\gamma(t)} + v(t)$ where $G_{\gamma(t)}(v(t)) = 1 \ \forall t$.

The condition that $G_{\gamma(t)}(v(t)) = 1$ and $G_{v(t)} < 1$ ensures that the navigator can always overcome the wind and progress can be made in every direction in T_pM at every point p on M.

2.2.2 General Solution Method

The Zermelo navigation problem has an elegant solution method which can be found in section (3) of [29] and many other places. This solution method involves finding another Finsler metric F on M for which the geodesics are the time optimal trajectories.

The equation that F must solve (point wise) is:

$$G_p\left(\frac{v}{F_p(v)} - W_p\right) = 1 \quad \forall v \in T_pM \tag{2.4}$$

One can show that the solution F is unique and always Finsler metric. Unfortunately, for most values of G, this equation cannot be solved in closed form.

2.2.3 Shen's Theorem

Initially Zermelo solved the navigation problem in the plane with time dependent wind. A full, slightly simpler, presentation of this solution can be found in [31]. The solution presented here, Shen's theorem, only applies to the time independent case but on a general manifold.

There is one case for which 2.4 can be solved for F in closed form. This is the case that G is, point wise, the norm induced by a Riemannian metric. That is to say $G_p(v) = \sqrt{h_p(v,v)}$ for some Riemannian metric h. A necessary and sufficient condition for this is that G_p satisfies the 'parallelogram identity' at each point p. This guarantees that a suitable h could be constructed by the standard procedure of the polarisation of a form on a vector space [32]. In this case the metric F can be obtained in closed form in terms of the navigation data (h, W). This was first achieved in remark (3.3) in [29] and is known as Shen's theorem.

The following derivation is after [29] and is known as Shen's theorem. In the following proof all point indicies p are dropped to reduce clutter. All equations are understood to hold point wise on M. In the Riemannian case, 2.4 reads:

$$\sqrt{h\left(\frac{v}{F(v)} - W, \frac{v}{F(v)} - W\right)} = 1 \tag{2.5}$$

The following derivation of a quadratic equation satisfied by F can now be made:

$$\begin{aligned}
\frac{1}{F^2(v)}h(v,v) - \frac{2}{F(v)}h(v,W) + h(w,w) &= 1 && \Rightarrow \\
F^2(v)(h(w,w) - 1) - 2F(v)h(v,W) + h(v,v) &= 0 && \Rightarrow \\
F^2(v) + \frac{2F(v)h(v,W)}{1 - h(w,w)} - \frac{h(v,v)}{1 - h(w,w)} &= 0 && \Rightarrow \\
F^2(v) + \frac{2F(v)h(v,W)}{\lambda} - \frac{h(v,v)}{\lambda} &= 0
\end{aligned} \tag{2.6}$$

where $\lambda := 1 - h(w,w)$ is a scalar function on M.

This equation can be solved using the standard quadratic formula to yield the following formulae defining F:

$$F_p(v) = -\frac{h_p(v, W_p)}{1 - h_p(W_p, W_p)} \pm \frac{\sqrt{h_p(v, W_p)^2 + \Big(1 - h_p(W_p, W_p)\Big)\, h_p(v, v)}}{1 - h_p(W_p, W_p)}$$

wherein the $\pm$ is chosen to ensure positivity of F. This can be rearranged into the following illuminating form:

$$\begin{aligned} F_p(v) &= \sqrt{\alpha_p(v, v)} + \beta_p(v) \\ \alpha_p(u, v) &= \frac{\lambda_p h_p(u, v) + \beta_p(u)\beta_p(v)}{\lambda_p^2} \\ \beta_p(v) &= -\frac{h_p(v, W_p)}{\lambda_p} \\ \lambda_p &= 1 - h_p(W_p, W_p) \end{aligned} \tag{2.7}$$

α as given above in 2.7 is a new Riemannian metric on M and $\beta \in \Gamma(T_p^*M)$ is a differential one form. One now concludes that Randers metrics solve the Zermelo navigation problem for Riemannian manifolds. A good review of Zermelo navigation on Riemannian manifolds can be found in [33].

2.3 Formulas in Coordinates

In coordinates the Riemannian metric α and the one form β appearing in the solution to Shen's theorem reads:

$$\begin{aligned} \alpha_{ij} &= \frac{h_{ij}}{\lambda} + \frac{W_i W_j}{\lambda^2} \\ \beta_j &= \frac{-W_i}{\lambda} \end{aligned} \tag{2.8}$$

where the lowering of the index on the vector field W has been done using the musical isomorphism of h i.e. $W_i = h_{ij}W^j$. In coordinates λ reads:

$$\lambda = 1 - h_{ij}W^iW^j \tag{2.9}$$

2.3.1 Inverting Shen's Solution

It is further possible to 'invert' Shen's solution 2.7:

$$\begin{aligned} h_{ij} &= \lambda\left(\alpha_{ij} - \beta_i\beta_j\right) \\ W^k &= -\frac{\beta^k}{\lambda} \end{aligned} \tag{2.10}$$

This establishes that every Randers metric can be expressed as the solution to a Zermelo navigation problem on a Riemannian manifold. In this sense the study of Randers metrics is sufficient for the study of Zermelo navigation on Riemannian manifolds. For the details of this any many other relevant facts about Randers metrics see chapters 1 and 2 of [23].

Chapter 3

Geometric Quantum Mechanics

Geometric quantum mechanics (GQM henceforth) is a reformulation of the standard theory of quantum mechanics in terms of differential geometry. Differential geometry enters the theory of QM in at least two ways, both the space of states and the space of time evolutions are differential manifolds. An excellent review of geometric quantum mechanics can be found here: [34].

Geometric quantum mechanics is being reviewed here in order to set up questions about the quantum speed limit in terms a Zermelo navigation problem.

3.1 Quantum Time Evolution, $SU(n)$ and $\mathfrak{su}(n)$

3.1.1 Quantum Time Evolution

In quantum mechanics the state of an isolated quantum system is typically described as a vector in $\left(\mathbb{C}^N, \langle\cdot|\cdot\rangle\right)$ (standard inner product) or an infinite dimensional Hilbert space.

The typical notation for quantum mechanics is the so called 'bra-ket' notation, any good quantum mechanics textbook covers this material [35, 36]. This is the choice of notation best adapted for present purposes and it is employed throughout this work.

In order to define quantum time evolution one first needs the following definition:

Definition 3.1.1 (Unitary Map/Operator)**.** A map $\hat{U} : \mathbb{C}^N \to \mathbb{C}^N$ is said to be *Unitary* if $\hat{U}^{-1} = \hat{U}^\dagger$.

This property of an operator is exactly the condition for an operator to preserve the length of all vectors in $\mathbb{C}^N$. This is because $\hat{U}|\psi\rangle$ has the property $\left(\hat{U}|\psi\rangle\right)^\dagger \hat{U}|\psi\rangle = \langle\psi|\hat{U}^\dagger\hat{U}|\psi\rangle = \langle\psi|\hat{U}^{-1}\hat{U}|\psi\rangle = \langle\psi|\psi\rangle$ for all $|\psi\rangle$.

The time evolution operator $\hat{U}_t$ is a unitary linear map from $\mathbb{C}^N \to \mathbb{C}^N$. The map $\hat{U}_t$ represents the time evolution of every pure state $|\psi\rangle$ via:

$$|\psi_t\rangle = \hat{U}_t|\psi_0\rangle \tag{3.1}$$

The standard formulation of QM time evolution, for closed systems with pure states, is the Schrödinger equation for $\hat{U}_t$:

$$\frac{d\hat{U}_t}{dt} = -i\hat{H}_t\hat{U}_t \tag{3.2}$$

As after no time, no change in any state can occur one sets: $\hat{U}_0 := \hat{I}$. One readily checks that $|\psi_t\rangle$ defined as in 3.1 satisfies the Schrödinger equation for the state:

$$\frac{d|\psi_t\rangle}{dt} = -i\hat{H}_t|\psi_t\rangle \tag{3.3}$$

when $\hat{U}_t$ satisfies 3.2.

3.1.2 The Role of $SU(n)$

The set of all possible time evolution operators is the unitary group, which is a Lie group.

Definition 3.1.2. A *Lie group* is a smooth manifold which is also a group for which the group multiplication is a smooth function.

For a full discussion of Lie groups see [37, 21]. Formally, the unitary group $U(n)$ is the group of linear operators mapping from $\mathbb{C}^n$ to $\mathbb{C}^n$ which preserve the standard L^2 inner product.

The group $U(n)$ can be understood as a subgroup of $GL(n)$, the group of all non singluar square complex matrices. This subgroup can be defined as:

$$\begin{aligned} U(n) &= \{\hat{U} \in GL(n) \text{ s.t } \hat{U}^\dagger = \hat{U}^{-1}\} \\ &= \{\hat{U} \in GL(n) \text{ s.t } \left|det\left(\hat{U}\right)\right| = 1\} \end{aligned} \tag{3.4}$$

The special unitary group, $SU(n)$ is a subgroup of $U(n)$. This is defined by:

$$SU(n) = \{\hat{U} \in U(n) \text{ s.t } det\left(\hat{U}\right) = 1\} \tag{3.5}$$

Both $U(n)$ and $SU(n)$ are linear algebraic groups. That is to say they are subsets of $GL(n)$ defined as the solutions to a set of polynomial equations. For the details of linear algebraic groups, see [38]. Both groups are compact and connected [39]. The dimensions of these two groups are:

$$\begin{aligned} \dim\left(U(n)\right) &= n^2 \\ \dim\left(SU(n)\right) &= n^2 - 1 \end{aligned} \tag{3.6}$$

Changes in global phases of a quantum state is not physically meaningful or measurable quantities even in principle. This is to say the transformation:

$$|\psi\rangle \mapsto e^{i\theta}|\psi\rangle \tag{3.7}$$

represents no physical change in a system in state $|\psi\rangle$. This can be readily checked, at least in the context of the standard POVM (positive-operator valued measure [40]) formalism for quantum measurements, by substituting the re-phased state into the formula for the probability of each possible measurement outcome and observing that the probability formula remains unchanged. For the details of POVMs and measurement theory see [41]. In the simplest description of a measurement, a Von Neumann measurement, this can be most easily confirmed. The probability of observing measurement outcome O_n, in state $|\psi\rangle$ when measuring some observable $\hat{O}$ is:

$$P_n = |\langle O_n|\psi\rangle|^2 \tag{3.8}$$

where $|O_n\rangle$ is the eigenstate of $\hat{O}$ associated to the eigenvalue O_n. Thus under the re-phasing transformation of the state:

$$P_n = |\langle O_n|\psi\rangle|^2 \mapsto |\langle O_n|e^{i\theta}|\psi\rangle|^2 = |e^{i\theta}|^2|\langle O_n|\psi\rangle|^2 = |\langle O_n|\psi\rangle|^2 = P_n \tag{3.9}$$

In light of this observation and after comparison to 3.1 one sees that the overall phase of $\hat{U}_t$ is also not physically relevant. This is to say:

$$\hat{U}_t \mapsto e^{i\theta(t)}\hat{U}_t \tag{3.10}$$

represents no change in the physical time evolution being represented. As such, we can choose $\theta(t)$ at our convenience. The choice made throughout this thesis is the value of $\theta(t)$ which renders $\hat{U}_t$ special unitary. This is:

$$\hat{U}_t \mapsto \left(\det(\hat{U}_t)\right)^{-M}\hat{U}_t \tag{3.11}$$

where M is the dimension of the Hilbert space on which $\hat{U}_t$ acts. It is noteworthy that this construction only makes sense in the case of finite dimension systems. One now has:

$$\begin{aligned}\det\left(\hat{U}_t\right) &\mapsto \det\left(\left(\det(\hat{U}_t)\right)^{-M}\hat{U}_t\right) \\ &= \left(\left(\det(\hat{U}_t)\right)^{-\frac{1}{M}}\right)^{M}\det\left(\hat{U}_t\right) \\ &= \det\left(\hat{U}_t\right)^{-1}\det\left(\hat{U}_t\right) = 1\end{aligned} \tag{3.12}$$

It is thus possible to only consider quantum time evolutions $\hat{U}_t \in SU(n)$ without any loss of physical content to the theory of quantum mechanics.

The advantage of this is that, as mentioned above, the dimension of the manifold underlying $SU(n)$ is lower than that of $U(n)$. The is also a further advantage, $SU(n)$ is a simple Lie group, unlike $U(n)$ where the intended definition is:

Definition 3.1.3. A *Simple Lie group* is a connected non-abelian Lie group which does not have nontrivial connected normal subgroups

Being simple is required to make the Killing form the unique, upto a constant multiple, invariant 2-form [21] on $\mathfrak{su}(n)$, a fact which will be used later.

3.2 $\mathbb{C}P^N$ As a Quantum State Space

In geometric quantum mechanics, the space of physically distinct states is described by $\mathbb{C}P^N$ a manifold. Each point of this manifold corresponds to a truly physically distinct state.

This material is being reviewed in order to later set up the problem of time optimal state transfer in geometric terms. The explanations in this section mostly closely follows [1].

3.2.1 Definition of $\mathbb{C}P^N$

Imposing Normalisation

The first step in realising CP^N as the space of physically distinct states is treating the normalisation condition. Define an equivalence relation $\sim_1 \subseteq \mathbb{C}^{N+1} \times \mathbb{C}^{N+1}$ by: $|\psi\rangle \sim_1 \lambda|\psi\rangle \ \forall\lambda \in \mathbb{R}/\{0\}$. This yields a sphere S^{2N+1} as a set of equivalence classes of points in $\mathbb{C}^{N+1}$. Each equivalence class consists of a (real) line through the origin in $\mathbb{C}^{N+1} \cong \mathbb{R}^{2(N+1)}$. This construction can also be clarified by writing an arbitrary state:

$$|\psi\rangle = \sum_k \alpha_k |A_k\rangle \tag{3.13}$$

where $\{|A_k\rangle\}$ is any orthonormal basis of $\mathbb{C}^{N+1}$. The condition that the state has norm one becomes:

$$\langle\psi|\psi\rangle = \sum_k |\alpha_k|^2 = \sum_k \Re(\alpha_k)^2 + \Im(\alpha_k)^2 = 1 \tag{3.14}$$

which is the formula defining a sphere of radius 1 as a subset of $\mathbb{R}^{2(N+1)}$.

Phase

The next step is to identify states which differ only by a global phase $e^{i\theta}$. This boils down to taking a quotient of S^{2N+1} into equivalence classes, where the equivalence relation $\sim_2 \subseteq S^{2N+1} \times S^{2N+1}$ is defined by $[|\psi\rangle]_{\sim_1} \sim_2 e^{i\theta}[\ |\psi\rangle\ ]_{\sim_1} \ \forall\theta \in \mathbb{R}$. The new space is:

$$S^{2N+1}/U(1) \cong \mathbb{C}P^N \tag{3.15}$$

This quotienting process can be represented by:

$$\mathbb{C}^{N+1} \xrightarrow{\phi_1} S^{2N+1} \xrightarrow{\phi_2} \mathbb{C}P^N, \qquad \mathbb{C}^{N+1} \xrightarrow{\gamma := \phi_2 \circ \phi_1} \mathbb{C}P^N$$

where:

$$\phi_1 : |\psi\rangle \mapsto [|\psi\rangle]_{\sim_1} \tag{3.16}$$

$$\phi_2 : [|\psi\rangle]_{\sim_1} \mapsto [[|\psi\rangle]_{\sim_1}]_{\sim_2} \tag{3.17}$$

Here $\gamma : \mathbb{C}^{N+1} \to \mathbb{C}P^N$ and realises the quotient by the equivalence relation $\sim \subseteq \mathbb{C}^{N+1} \times \mathbb{C}^{N+1}$ defined by $|\psi\rangle \sim \lambda e^{i\theta}|\psi\rangle \; \forall \lambda \in \mathbb{R}/\{0\}, \theta \in \mathbb{R}$. The purpose of breaking the quotienting process into two steps is to illustrate the mathematical realisation of the two physical principles: that states are normalised and that global phases are unphysical. This equivalence is the same as $|\psi\rangle \sim Z|\psi\rangle \; Z \in \mathbb{C}/\{0\}$. Thus one sees that each equivalence class is a complex line.

Henceforth where the 'equivalence class $[|\psi\rangle]$ of a state $|\psi\rangle$' is referred to, this is taken to mean the corresponding single point in $\mathbb{C}P^n$ given by $\gamma(|\psi\rangle)$.

3.2.2 A Common Misunderstanding

It is often misstated in physics literature that $U(n) \cong SU(n) \times U(1)$. This statement requires clarification. It is true if the correct meaning of the $\times$ symbol is understood. The statement holds true as stated in terms of the topological and smooth structures of $U(n), SU(n)$ and $U(1)$. This is to say that $U(n)$ is the *manifold product* (for details see [42]) of $SU(n)$ and $U(1)$.

However, the same statement interpreted with $\times$ referring to the direct product of groups, is false. This can be directly shown to be false by applying the well known first isomorphism theorem for groups to the homomorphism $\phi : SU(N) \times U(1) \to U(N)$ given by $\phi(\hat{U}, e^{i\theta}) = e^{i\theta}\hat{U}$. The kernel of the homomorphism ϕ is the set $\{e^{i\theta}\hat{U} \text{ s.t } e^{i\theta} \in \mathbb{C} \text{ is an } N^{th} \text{ root of unity}\}$, which as a multiplicative group is $\mathbb{Z}_N$, the integers mod N.

The correct statement is instead that $U(N)$ can be written as a semi-direct product of $SU(N)$ and a $U(1)$ subgroup of $U(N)$. The object which behaves as $U(N)$ with operators identified up to a global phase is the projective unitary group $PSU(N)$. It is defined by the quotient $U(N)/Z(U(N))$ where $Z(U(N))$ is the centre of $U(N)$, well know to be the 'scalar matrices' $\{e^{i\theta}\hat{I}|\theta \in [0, 2\pi)\}$. It is true, in the sense of the product of groups, that $U(N) \cong (SU(N) \times U(1))/\mathbb{Z}_N$, which is the conclusion of applying the first isomorphism theorem to the homomorphism ϕ.

The Lie algebras of $PSU(N)$ and $SU(N)$ are identical, working with $SU(N)$ rather than $PSU(N)$ does not affect anything that follows in this thesis. As such, it is $SU(n)$ which is used as the group of quantum time evolutions in the remainder of this thesis.

3.2.3 $\mathbb{C}P^n$ As a Homogeneous Space

A homogeneous space is a manifold M with the smooth, transitive action $\phi : G \times M \to M$ of a Lie group G defined on it [43]. In such a scenario one can show the existence of the following isomorphism:

$$M \cong G/\text{Stab}(p) \tag{3.18}$$

for any $p \in M$. Here the stabiliser subgroup of p is defined as:

$$\text{Stab}(p) = \{p \in M \text{ s.t. } \phi(p) = p\} \tag{3.19}$$

It is easily checked that, if G and M are compact then so is $\text{stab}(p)\ \forall p \in M$.

As $SU(n+1)$ acts transitively on $\mathbb{C}P^n$, the mathematical relationship between the group of all time evolutions $SU(n)$ and the space of states $\mathbb{C}P^n$ possess a special geometrical relationship:

$$\mathbb{C}P^N \cong SU(N+1)/U(N) \tag{3.20}$$

Care must be taken with this definition as, as it currently stands, it is ambiguous. In order to see this, firstly, the definition of $SU(N+1)/U(N)$ must be given. The quotient appearing in 3.20 is similar to a quotient group in the sense that $SU(N+1)/U(N)$ represents the set of right cosets of $SU(N+1)/U(N)$:

$$SU(N+1)/U(N) = \left\{\hat{U}U(N) \text{ s.t. } \hat{U} \in SU(n+1)\right\} \tag{3.21}$$

It must be noted that this quotient does not define a new group as $U(n)$ is not a *normal* subgroup of $SU(n+1)$ which is required to define an unambiguous rule for multiplying cosets.

Any $U(n)$ subgroup constructed as the stabiliser subgroup of some point $\hat{U} \in SU(n)$ of $SU(n)$ can be used, the resulting construction is isomorphic (diffeomorphic). A particularly simple choice is to take the point for which $U(n)$ is the stabiliser to be the equivalence class of the vector:

$$|\psi_0\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{3.22}$$

One can readily check that the group of special unitary matrices which preserve the equivalence class of this vector is:

$$\text{stab}\left([|\psi_0\rangle]\right) = \left\{ \begin{pmatrix} \det(\hat{V})^{-1} & 0\cdots 0 \\ 0 & \\ \vdots & \hat{V} \\ 0 & \end{pmatrix} \text{ s.t. } \hat{V} \in U(N) \right\} \cong U(N) \tag{3.23}$$

On a very technical note, this resulting space has a unique smooth structure compatible with the quotient map $\pi : SU(n+1) \to SU(n+1)/U(n)$. This map is defined by $\pi(\hat{U}) = \hat{U}U(n)$ and the unique smooth structure is the determined by insisting that this map is smooth. For the details of this see [43].

It is important to notice what has and has has not been achieved by this construction. This is an equivalence of the spaces $\mathbb{C}P^n$ and $SU(n+1)/U(n)$ as smooth manifolds. No equivalence of algebraic structures is implied despite the fact that the group operations of $SU(n+1)$ and $U(n)$ have been refereed to in the construction. There is no group structure on $\mathbb{C}P^n$.

Chapter 4

Speed Limits In Driven Quantum Systems

There has been much recent interest in driven quantum systems. Recent work on time optimal quantum gates implemented in a variety of quantum systems includes [44, 45, 46, 22, 47, 4]:

- [44] discusses time optimal implementation of a number of two qubit gates and also discuss experimental implementations of such gates.
- Work on open-dissipative systems for implementing quantum gates can be found in [45].
- Some works on this topic based on geometric methods include [46, 22].
- [46] discusses the use of sub-Riemannian metrics on the unitary group with application to two and three qubit systems, special focus on NMR experiments is given.
- [22] analyses the use of metric structure (in the sense of metric spaces, not differential geometry) to determining the quantum speed limit (QSL hensforth) for implementing quantum gates.
- [47] connects the QSL for orthogonality times and the QSL for implementing quantum gates.
- [4] produces a result based on a variational principle for a Lagrangian on $U(N)$, this work also shows how optimal control schemes can be obtained via differential geometry.

4.1 Driven Quantum Systems

In quantum optimal control, a certain type of system is ubiquitous. These are driven systems with a 'drift term'. Such systems are also ubiquitous within control theory more generally [48]. In this thesis, only systems with finite dimensional Hilbert spaces and pure states are discussed unless explicitly stated. The term 'drift' is chosen as this term represents the system's dynamics in the absence of any control. An example of a 'drift' in control theory include, the dynamics of a boat on a windy sea when the motors are all off and this is the origin of the term. Such systems take a specific mathematical form.

The Schrödinger equation for $\hat{U}_t$, the time evolution operator [49], standardly reads:

$$\frac{d\hat{U}_t}{dt} = -i\hat{H}_t\hat{U}_t \tag{4.1}$$

A common scenario is: $\hat{H}_t = \hat{H}_0 + \hat{H}_c(t)$ which yields:

$$\frac{d\hat{U}_t}{dt} = -i(\hat{H}_0 + \hat{H}_c(t))\hat{U}_t \tag{4.2}$$

for the equation governing the dynamics of $\hat{U}_t$. $\hat{H}_c(t)$ is known as the control Hamiltonian, it represents the effects of external control fields on the system. The interpretation is typically that the 'drift' term $\hat{H}_0$ is constant in time and outside the control of experimenters, at least during the time evolution of $\hat{U}_t$. This term typically represents the properties of a system, the state of which is being controlled. Systems of the form 4.2 are known as affine control systems. All the systems in this thesis have Hermitian ($\hat{H}^\dagger = \hat{H}$) Hamiltonians. Other systems, in PT symmetric quantum mechanics, have been considered in the context of the QSL [50].

The group in the case of quantum dynamics of finite dimensional systems is typically taken to be $U(N)$, the unitary group. However, we specialise throughout to $SU(N)$. This is easily achieved mathematically (by considering only traceless Hamiltonians) and is without effect on any of the physical predictions of the theory.

The general solution to the Schrödinger equation 4.1 for $\hat{U}_t$ can be expressed using a variety of series expansions. Notable among these are the *Dyson Series* [51], the *Magnus Expansion* [52] and the Fer infinite product expansion [53]. While the Magnus expansion has favorable mathematical properties at every order of approximation to the solution (all orders are unitary), the Dyson series is arguably easier to work with. The Dyson series reads:

$$\hat{U}_t = \mathcal{T}\exp\left(-i\int_0^t \hat{H}_{t'}dt'\right) \tag{4.3}$$

Where the $\mathcal{T}$ indicates time ordering the terms in the Taylor expansion of the matrix exponential. For an initial segment of the series see [51] or any good text on quantum field theory as this series is commonly applied in that context.

One now readily checks that the transformation $f : \mathfrak{u}(n) \to \mathfrak{su}(n)$ which removes the trace of the Hamiltonian:

$$f : i\hat{H}_t \mapsto i\hat{H}_t - \frac{\mathrm{Tr}(i\hat{H})}{n}\hat{I} \tag{4.4}$$

results:

$$\hat{U}_t = \mathcal{T}\exp\left(-i\int_0^t \hat{H}_{t'}dt'\right) \mapsto \mathcal{T}\exp\left(-i\int_0^t \hat{H}_{t'} - \frac{\mathrm{Tr}(\hat{H}_t)}{n}\hat{I}dt'\right) \tag{4.5}$$
$$= \mathcal{T}\exp\left(-i\int_0^t \hat{H}_{t'}dt'\right)\exp\left(-\frac{i}{n}\int_0^t \mathrm{Tr}(\hat{H}_{t'})dt'\right)$$
$$= \exp\left(-\frac{i}{n}\int_0^t \mathrm{Tr}(\hat{H}_{t'})dt'\right)\hat{U}_t$$

As $\exp\left(-\frac{i}{n}\int_0^t \mathrm{Tr}(\hat{H}_{t'})dt'\right)$ is simply a unit complex number, all that has happened is that $\hat{U}_t$ has acquired a phase factor. This clearly does not affect any physical predictions as is standard and well known.

In light of this we always restrict to $\hat{U}_t \in SU(n)$ by imposing that $\mathrm{Tr}(\hat{H}_t) = 0$ for all systems under consideration; as we have seen, this is entirely without loss of physical generality.

4.1.1 Affine Control Systems on $SU(n)$

The system 4.2 commonly takes a particular form in quantum control problems.

$$\frac{d\hat{U}_t}{dt} = -i\left(\hat{H}_0 + \sum_{k=0}^{K} f_k(t)\hat{H}_k\right)\hat{U}_t \tag{4.6}$$

Where K is a, typically small, integer. In such a scenario, the f_k are referred to as the control fields. Each $\hat{H}_k$ represents the effect of the kth control field on the system evolution.

4.2 Work of Carlini Et Al

This section reviews two recent papers: [5] and [4].

- Firstly, [5] studies the question: how can a given initial state $|\psi_0\rangle$ be transformed into a given terminal state $|\psi_1\rangle$ in a time optimal way.
- Secondly, [4] studies the question: how can $\hat{H}_t$ be chosen, subject to constraints, so that a given gate $\hat{O} \in SU(n)$ is implemented in a time optimal way.

Throughout this thesis, these two questions are referred to as the first and second fundamental problems respectively.

4.2.1 'The Quantum Brachistochrone'

The paper [5] is concerned with time optimally transforming the initial pure state $|\psi_0\rangle$ of a finite dimensional quantum system into a desired terminal state $|\psi_1\rangle$ in the least time possible in the presence of constraints. The classical brachistochrone is not reviewed in this thesis as it has been described excellently in many places [54].

In this thesis the problem of time optimal quantum evolution is posed for systems with finite dimensional states space $\mathbb{C}^{n+1}$ and pure states. The, time dependent, *overall* (in contrast to the work in 6 where only the control is constrained) Hamiltonian $\hat{H}_t$ is constrained such that $\mathrm{Tr}(\tilde{H}^2) = 2\omega^2$. Here $\tilde{H}$ is the traceless part of the Hamiltonian, $\tilde{H} = \hat{H} - \frac{\mathrm{Tr}(\hat{H})}{N}\hat{I}$. This can be mathematically understood as expressed earlier 4.4. The time optimal trajectories of $|\psi_t\rangle$ are determined by a variational method for an action defined on $\mathbb{C}P^n$.

Later in the paper a more general set of constraints is considered. A set of M functions $\{f_k : \frac{1}{i}\mathfrak{su}(n) \to \mathbb{R} | k = 0 \dots M\}$ is given and the constraint that $f_k(\hat{H}_t) = c_k$ (where each c_k is a constant) is imposed for all time during an evolution. Here only the more general methodology is reviewed.

The action considered in the variational principle is defined as:

$$S\left[|\psi_t\rangle, \hat{H}_t, |\phi_t\rangle, \lambda_t\right] = \int_0^T \frac{\sqrt{\langle\delta\psi_t|\left(\hat{I} - P_t\right)|\delta\psi_t\rangle}}{\Delta E_t} + \left(i\langle\delta\phi_t|\psi_t\rangle - i\langle\phi_t|\delta\psi_t\rangle + 2\langle\phi_t|\hat{H}_t|\psi_t\rangle\right) dt + \sum_{k=0}^{M} \lambda^k f_k(\hat{H}_t) \tag{4.7}$$

where:

- $|\psi_t\rangle$ is the system's state
- $\hat{H}_t$ is the system's Hamiltonian
- $\langle\phi_t|$ is a vector Lagrange multiplier in $\mathbb{C}^{n+1*}$, the dual of the state space
- λ^k are scalar Lagrange multipliers, which can be taken to be constants by the argument given in the original paper which is not reproduced here
- $P_t := |\psi_t\rangle\langle\psi_t|$, this is mathematically a projection operator onto $|\psi_t\rangle$. Physically it is the density matrix of $|\psi_t\rangle$.

Within the paper it is argued that the time optimal trajectories are the minimising curves for this action. The variations considered are of $|\psi_t\rangle$, $\hat{H}_t$ and variation by all the Lagrange multipliers.

The paper goes on to define an operator F by:

$$\hat{F} := \sum_k \lambda^k \frac{\delta f_k}{\delta \hat{H}} - \left\langle \frac{\delta f_k}{\delta \hat{H}} \right\rangle \hat{P}_t \tag{4.8}$$

wherein the expectation brackets $\left\langle \frac{\delta f_k}{\delta \hat{H}} \right\rangle$ indicate the expectation of the operator $\frac{\delta f_k}{\delta \hat{H}}$ in the state $|\psi_t\rangle$, i.e. $\langle\psi_t|\frac{\delta f_k}{\delta \hat{H}}|\psi_t\rangle$. It is then shown that the traceless part of the optimal Hamiltonian, $\tilde{\hat{H}}$, satisfies:

$$\left(\dot{\hat{F}} + i\left[\tilde{\hat{H}}, \hat{F}\right]\right)|\psi_t\rangle = 0 \tag{4.9}$$

This equation is solved in closed form for two simple cases for a single spin.

Spatially Isotropic Constraint

The constraint treated is:

- $f(\hat{H}_t) = \frac{1}{2}\operatorname{Tr}\left(\tilde{\hat{H}}_t^2\right) - \omega^2 = 0$ for some real number ω for all time

The optimal trajectories for the state are shown to be the Fubini-Study geodesics on $\mathbb{C}P^N$. The optimal Hamiltonian (taken to be traceless) driving from $|\psi_\mathrm{i}\rangle$ to $|\psi_\mathrm{f}\rangle$ are shown to be:

$$\hat{H} = i\omega\left(|\psi_\mathrm{f}\rangle\langle\psi_\mathrm{i}| - |\psi_\mathrm{i}\rangle\langle\psi_\mathrm{f}|\right) \tag{4.10}$$

which does not depend on time.

It is also shown the the optimal time for this transition with the given constraint is:

$$T^* = \frac{1}{|\omega|}\arccos\left(|\langle\psi_\mathrm{i}|\psi_\mathrm{f}\rangle|\right) \tag{4.11}$$

Spatially Anisotropic Constraint

The constraints treated are:

- $f_1(\hat{H}_t) = \frac{1}{2}\operatorname{Tr}\left(\tilde{\hat{H}}_t^2\right) - \omega^2 = 0$ for some real number ω for all time
- $f_2(\hat{H}_t) = \operatorname{Tr}(\tilde{\hat{H}}_t\ \sigma_z) = 0$ for all time

This imposes that $\operatorname{Tr}\left(\tilde{\hat{H}}^2\right) = 2\omega$ for all time. In this case 4.9, which is eqn. (16) in the original work, can be solved to find:

$$\begin{aligned}\hat{U}_t &= \exp\left(it\Omega\sigma_z\right)\exp\left(-it\left(\tilde{\hat{H}}(0) + \Omega\sigma_z\right)\right) \\ \tilde{\hat{H}}_t &= \exp(it\Omega\sigma_z)\tilde{\hat{H}}(0)\exp(-it\Omega\sigma_z)\end{aligned} \tag{4.12}$$

wherein Ω can be expressed as the ratio of the two Lagrange multipliers appearing in the Lagrangian. These are eqns (22) and (23) in the original work.

4.2.2 'Time Optimal Unitary Operators'

The paper [5] treats the question of obtaining time optimal Hamiltonians, subject to constraints, for implementing a quantum gate $\hat{O}$. This question is posed by asking which trajectory (consistent with the constraints) the time evolution operator $\hat{U}_t$ must follow in order to connect the identity $\hat{I}$ to the desired gate $\hat{O}$ in the least time.

A variational principle on the unitary group $U(N)$ is introduced using the following action functional:

$$L[\hat{U}_t, \hat{H}_t, \hat{\Lambda}_t, \lambda_t] = \int_0^T \sqrt{\frac{\left\langle \frac{d\hat{U}_t}{dt}, \left(\hat{I} - P_{\hat{U}_t}\right)\left(\frac{d\hat{U}_t}{dt}\right)\right\rangle}{\left\langle \hat{H}_t\hat{U}_t, \left(\hat{I} - P_{\hat{U}_t}\right)\left(\hat{H}_t\hat{U}_t\right)\right\rangle}} + \left\langle \hat{\Lambda}_t, i\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger - \hat{H}_t \right\rangle dt \sum_{k=0}^{M} \lambda^k f_k(\hat{H}_t) \tag{4.13}$$

where:

- $\hat{U}_t \in U(N)$ is the time evolution operator, which is not a required to be *special* unitary a priori
- $\hat{H}_t$ is the Hamiltonian, which is taken to be Hermitian but not required to be traceless
- $\hat{\Lambda}_t$ is a Hermitian matrix, $\langle \hat{\Lambda}_t, \cdot \rangle$ (an element of the dual space to the vector space of all Hermitian matrices) is a Lagrange multiplier
- $P_{\hat{V}}$ is, for each $\hat{V} \in U(N)$, a map from complex matrices to complex matrices. This map is defined by $P_{\hat{V}}(\hat{A}) = \frac{1}{N}\operatorname{Tr}(\hat{A}\hat{U})\hat{V}$
- $\langle \cdot, \cdot \rangle$ is the Hilbert Schmidt inner product defined by $\langle \hat{A}, \hat{B} \rangle = \operatorname{Tr}(\hat{A}^\dagger \hat{B})$
- λ^k are scalar Lagrange multipliers, which can be taken to be constant (by a similar argument to that in [4]) which is not reproduced here

The paper goes on to define:

$$\hat{F} = \frac{\partial \hat{L}_c}{\partial \hat{H}_t} \tag{4.14}$$

and to show that the optimal overall Hamiltonian must satisfy:

$$i\frac{d\hat{F}}{dt} = [\hat{H}_t, \hat{F}] \tag{4.15}$$

by applying a variational argument to the Lagrangian 4.13. Some simple cases are solved in closed form. Specifically, the cases of two qubit gates for a system with anisotropic Heisenberg coupling are both solved. The optimal Hamiltonians are determined in the case of the *swap* and entangler gates.

4.2.3 Comments On The Work Of Carlini Et Al.

The work of Carlini Et Al. [4, 5] is highly effective and interesting. However, the method is limited in a specific practically relevant way.

The method yields an equation satisfied by the optimal Hamiltonian (eqn. (16) in [5]) driving the time evolution operator $\hat{U}_t$ along a stationary curve of the action functional 4.13. This work has provided a method of determining the time optimal trajectories for a quantum system with constraints. However, it does not provide a method for calculating the optimal time for $\hat{U}_t$ to traverse a given curve on $SU(N)$. This is a physically relevant requirement for practical use in quantum control applications as not all trajectories can be always be physically realised given only a restricted set of admissible control fields. This is the case even when a system is controllable.

The work [55] also studies the possibility of applying numerical methods to solve the eqn (16) in [4] with a good deal of success in the two qubit case. However, the required calculations are extremely involved. If a simpler way could be found then this would be favourable.

Chapter 5

The Physical Limits To Computation: Quantum Mechanics and Further Afield

Within this section, the context of the broader study of natural and non-standard computation are outlined in order to provide motivation for the later chapters.

This section is included in order to briefly cover related notions of the speed limit to computation. Not all the limits discussed are directly related to the in this thesis research and are included for completeness of context.

5.1 Natural Computation

5.1.1 Grand Challenge

In 2002, the UK Computing Research Committee (UKCRC) issued a series of grand challenges for computing research [56, 57]. The GC7 Challenge has been formally stated as:

> 'to produce a fully mature science of all forms of computation, that unifies the classical and non-classical paradigm'

This sets the challenge of putting the theory of computation on solid physical grounds and assessing the physical limitations to computation. It also sets the challenge of establishing these limits in a way comparable to the known classical limits in terms of Turing machine and computational complexity.

5.1.2 When Does a Physical System Implement a Computation?

There has been a great deal of discussion on the topic of when a physical system implements a specific computation from within different areas of science [58], computer science [59], mathematics [60] and philosophy [61] to name but only a few papers. The issue is far from settled. Although many varied proposals have been made for resolving this issue, one must be chosen if any progress is to be made.

Ultimately, an implicit operational assumption is routinely in use in quantum computation. This is that a system (with time evolution operator $\hat{U}_T$ at time T) implements a computation when $\hat{U}_T = \hat{O}$. Here $\hat{O}$ is a desired quantum 'gate'.

In the case of many other physical systems which can implement computations, this mapping from elementary computational operations to possible physical time evolutions implementing them is far less canonical. Interestingly the classical case is less developed than the quantum one.

5.2 Non-Relativistic Quantum Speed Limits

5.2.1 Lieb-Robinson Bound

The Lieb-Robinson bound (LRB) is a bound on the rate of information transfer in non-relativistic quantum systems with a lattice structure [62]. For a full bibliography and history of this concept see [63]. The bound is expressed in terms of the maximum speed a wave can propagate in a lattice of spins in order to carry information.

This bound is mentioned for completeness and is not discussed further.

5.2.2 Margolis Levitin Theorem and Mandelshtam-Tamm inequality

The Margolis-Levitin bound [64] (ML bound) is a well known speed limit to QIP.

The Margolis-Levitin Theorem provides a bound on the rate of dynamical evolution of a (non-relativistic) quantum system in terms of its total energy [64]. In order to understand the motivation for interpreting the time taken to transition between orthogonal states of a quantum system as rate of computation one must first discuss the nature of distinguishable states.

Which States Can Be Distinguished with Certainty By a Single Measurement?

Consider a quantum system that is known with certainty to reside in one of two (normalised) quantum states $|\psi\rangle$ or $|\psi\rangle$. Consider the observable formed as the outer product $\hat{O} := |\psi\rangle\langle\psi|$, i.e. a projection operator (as it is clearly idempotent) on to the state $|\psi\rangle$. The eigenstates and eigenvalues of this operator (which will be interpreted, as usual for a projective measurement, as measurement outcomes and post-measurement states respectively) are simple to find and can be obtained in the following standard manner:

$$
\begin{aligned}
\hat{O}|\psi\rangle &= (|\psi\rangle\langle\psi|)\,|\psi\rangle = |\psi\rangle\langle\psi|\psi\rangle = 1|\psi\rangle \\
\hat{O}|\phi\rangle &= (|\psi\rangle\langle\psi|)\,|\phi\rangle = |\phi\rangle\langle\phi|\psi\rangle = 0|\phi\rangle
\end{aligned}
\tag{5.1}
$$

This shows that both $|\psi\rangle$ and $|\phi\rangle$ are eigenstates of $\hat{O}$ with eigenvalues 1 and 0 respectively. One can now conclude the contents of the following table 5.2.2 of outcome probabilities:

	State Before Measurement	
Outcome	$\|\psi\rangle$	$\|\phi\rangle$
0	0	1
1	1	0

This shows that in fact these two states can in principle distinguished with certainty by measuring the observable $\hat{O}$. Orthogonality is both necessary and sufficient for the existence of an observable which can be used to distinguish two states with certainty. For fuller detail about quantum measurement and metrology (specifically in the interesting context of control) see [65].

The speed of a classical computer can be considered to be the maximum number of states the computer can pass through during a computation per unit of time [64, 66]. The interpretation of the inequality given in [64] as a lower bound for the running time of a computation depends on the interpretation of distinguishable (orthogonal) quantum states as the appropriate corresponding concept. While this is not the only way to quantify the speed of a computer, it is one way which has attracted significant interest and follow-up [67, 68, 69, 70].

ML Bound

Consider a quantum system with Hamiltonian H (taken to be Hermitian as usual) and also consider its energy eigenstates (the energy basis) and energy eigenvalues. For simplicity [64] considers only Hamiltonians with discrete spectra. I denote $\hat{H}|\phi_i\rangle = E_i|\phi_i\rangle$ where $i \in (\mathbb{N} \cup 0)$ for the energy eigenstates and eigenvalues; this numbering of states is taken to be such that the eigenvalues are non-decreasing in i. Following [64], the ground state $|\phi_0\rangle$ (which is assumed to exist) is taken to have energy 0. This is always possible as potential energies are only defined up to an additive constant. As per the previous discussion of orthogonal states one may want to ask what the shortest possible time for any given $|\psi_0\rangle$ to evolve (under the time evolution of a given H) in to a state orthogonal to $|\psi_0\rangle$. The central result of [64] is that a bound can be placed on the orthogonality time $\tau_\perp$ in terms of the average energy of the state $|\psi_0\rangle$.

The novel bound given in [64] is:

$$\tau_\perp \geq \frac{\pi\hbar}{2E} \tag{5.2}$$

where E is the average energy of $|\psi_0\rangle$ (with $E_0 = 0$).

Assumptions Underlying the Margolis Levitin Theorem

The Margolis Levitin theorem has many physical and computational assumptions underlying it. Some of these include:

- It is a statement about non-relativistic systems only, a similar statement has been proposed for relativistic systems [70].
- The only measure of computation speed to which it applies is the orthogonality time. If certainty is not required, i.e. a probabilistic result to a computation is acceptable, then only considering distinguishable states may not be the best measure of computation speed.
- It says nothing about the number of *distinct* orthogonal states states a system can evolve into in a time interval.
- It makes no distinction concerning *which* operations a computer can perform, just how many operations it can perform (in terms of the minimal orthogonality time).
- It only applies to time independent Hamiltonians. Another similar bound which applies to to independent Hamiltonians does exist and can be found it [67].
- It only bounds the orthogonality time in terms of the expectation of the energy and ignores other physical properties of the system. This is both a strength, as it brings general applicability, and a weakness as some other physical property of the system may be the true limiting factor. For example, as the energy increases a system may become unstable and no longer remain spatially bounded.

Proof Technique

The following is directly after [64]. Here one notes that, by well known properties of separable Hilbert spaces and Hermitian operators on such spaces, any state $|\psi_0\rangle$ can now be written $|\psi_0\rangle = \sum_{i=0}^{i=\infty} c_i|\phi_i\rangle$ for some complex coefficients $\{c_i\}$. One also notes that the time evolution, the the systems to which the bound applies have time independent Hamiltonians, of $|\psi_0\rangle$ is given by $|\psi_t\rangle = e^{\frac{-itH}{\hbar}}|\psi_0\rangle$. This can be expanded in terms of the expansion $|\psi_0\rangle = \sum_{i=0}^{i=\infty} c_i|\phi_i\rangle$ and the expansion for the exponential of an operator $e^{\frac{-itH}{\hbar}} = \sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{-itH}{\hbar}\right)^k$ as follows:

$$|\psi_t\rangle = e^{\frac{-itH}{\hbar}}|\psi_0\rangle = \sum_{i=0}^{\infty} e^{\frac{-itE_i}{\hbar}} c_i|\phi_i\rangle \tag{5.3}$$

From this one obtains:

$$\begin{aligned} S(t) := \langle\psi_0|\psi_t\rangle &= \left(\sum_{i=0}^{\infty} \overline{c_i}\langle\phi_i|\right)\left(\sum_{j=0}^{\infty} e^{\frac{-itE_j}{\hbar}} c_j|\phi_j\rangle\right) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \overline{c_i}c_j e^{\frac{-itE_j}{\hbar}}\langle\phi_i|\phi_j\rangle \\ &= \sum_{i=0}^{\infty} |c_i|^2 e^{\frac{-itE_i}{\hbar}} \end{aligned} \tag{5.4}$$

where the final step has used the fact that $\langle\phi_i|\phi_j\rangle = \delta_{ij}$ which follows from the fact that $|\phi_i\rangle$ are eigenstates of a Hermitian operator. One seeks to determine the minimal t such that $S(t) = 0$. The solutions is obtained as follows:

$$\Re(S(t)) = \sum_{i=0}^{\infty} |c_i|^2 \Re\left(e^{\frac{-itE_i}{\hbar}}\right) = \sum_{i=0}^{\infty} |c_i|^2 \cos\left(\frac{tE_i}{\hbar}\right) \tag{5.5}$$

$$= 1 - \frac{2t}{\pi\hbar}\left(\sum_{i=0}^{\infty} |c_i|^2 E_i\right) + \Im(S(t)) \tag{5.6}$$

$$= 1 - \frac{2t}{\pi\hbar}\langle H\rangle_{\psi_0} + \Im(S(t))$$

$$= 1 - \frac{2t}{\pi\hbar}E + \Im(S(t))$$

If $S(t) = 0$ then $\Re(S(t)) = 0$ and $\Im(S(t)) = 0$. Noting that:

$$\forall x \in \mathbb{R},\ x \geq 0 \Rightarrow \cos(x) \geq 1 - \frac{2}{\pi}(x + \sin(x)) \tag{5.7}$$

it follows that:

$$\begin{aligned} S(t) &= 0 \Rightarrow \\ 0 \leq 1 &- \frac{2(\tau_\perp)}{\pi\hbar}E \Rightarrow \\ \tau_\perp &\geq \frac{\pi\hbar}{2E} \quad \square \end{aligned} \tag{5.8}$$

Mandelshtam-Tamm inequality

The ML bound contrasts with a previously known bound in terms of the uncertainty of the energy, $\Delta E := \Delta\hat{H}$ [71]. The Mandelstam-Tamm inequality is also a bound on the speed of computation [64], see [71] for an extensive review. Here $\Delta\hat{O}$ for any Hermitian operator $\hat{O}$ acting on the states of a system in state $|\psi\rangle$ is defined as $\Delta\hat{O} := \sqrt{\langle\psi|O^2|\psi\rangle + \langle\psi|O|\psi\rangle^2}$. It is also known that, with the same meaning for $\tau_\perp$:

$$\tau_\perp \geq \frac{\pi\hbar}{2\Delta E} \tag{5.9}$$

This statement has undergone a lot of analysis [71], a lot of misunderstanding [72] and many different derivations [71]. There is also a closely analogous statement concerning quantum field theory. It is worst noting that the is ongoing conflicting terminology in use in both the physics and mathematics community about which statment exactly is the 'time-energy uncertainty relation'.

Mandelshtam-Tamm inequality and Fubini-Study Geodesics

There is a unique Riemannian metric, up to a constant multiple, on $\mathbb{C}P^n$ (for all n) that is invariant under the natural choice of action of the unitary group on $\mathbb{C}P^n$. That is to say it has the property that $U(n)$ consists only of isometries [34]. This metric is the Fubini-Study metric; there is more than one way to represent its metric tensor.

One such way, well adapted for use in quantum mechanics, is to write it as a function on the tangent spaces to $\mathbb{C}^{n+1}$ that is constant on equivalence classes of the equivalence relation that allows $\mathbb{C}P^n$ to be constructed from $\mathbb{C}^{n+1}$ as described earlier. The formula is:

$$ds^2 = \frac{\langle \delta\psi|\delta\psi\rangle}{\langle \psi|\psi\rangle} - \frac{\langle \delta\psi|\psi\rangle\langle \psi|\delta\psi\rangle}{\langle \psi|\psi\rangle^2} \tag{5.10}$$

This metric can readily be used to prove the Mandelshtam-Tamm inequality by knowing that the geodesic distance between orthogonal states in Hilbert space is $\frac{\pi}{2}$. This can be observed by noting that the finite form of this metric is [73]:

$$\gamma(|\psi\rangle, |\phi\rangle) = \arccos\sqrt{\frac{|\langle \psi|\phi\rangle|^2}{\langle \psi|\psi\rangle^2\langle \phi|\phi\rangle^2}} \tag{5.11}$$

The metric clearly has the same unitary invariance properties as its infinitesimal form, 5.10.

The following relationship between the Mandelshtam-Tamm inequality and the Fubini-Study metric is well known; we re-derive it to illustrate the usefulness of geometric constructions in proving quantum speed limit theorems. In the case that $|\psi_t\rangle$ solves the Schrödinger equation for a time-independent Hamiltonian $\hat{H}$ we have:

$$|\delta\psi_t\rangle = \frac{d}{dt}|\psi_t\rangle = \frac{d}{dt}\exp\left(-it\hat{H}\right)|\psi_0\rangle = -i\hat{H}\exp(-it\hat{H})|\psi_0\rangle = -i\hat{H}|\psi_t\rangle \tag{5.12}$$

Substituting this into the definition of arc length corresponding to the metric at hand, we find:

$$\begin{aligned} L[|\psi_t\rangle] &= \int_{t=0}^{t=\tau} \sqrt{\langle \delta\psi_t|\delta\psi_t\rangle - \langle \delta\psi_t|\psi_t\rangle\langle \psi_t|\delta\psi_t\rangle}dt \\ &= \int_{t=0}^{t=\tau} \sqrt{\langle \psi_t|\hat{H}^2|\psi_t\rangle - \langle \psi_t|\hat{H}|\psi_t\rangle^2}dt \\ &= \int_{t=0}^{t=\tau} \Delta E_{|\psi_t\rangle}dt = \int_{t=0}^{t=\tau} \Delta E_{|\psi_0\rangle}dt = \tau\Delta E_{|\psi_0\rangle} \geq \frac{\pi}{2} \end{aligned} \tag{5.13}$$

$\Delta E_{|\psi_t\rangle}dt$ can be replaced by $\Delta E_{|\psi_0\rangle}$ in the last line, since here the Hamiltonian is time-independent, which implies that the energy uncertainty is also. From this follows the Mandelstam-Tamm inequality:

$$\tau \geq \frac{\pi}{2\Delta E_{|\psi_0\rangle}} \tag{5.14}$$

Compare this derivation to that in [74] (their eqns. 22-25; note that the 'Wootters distance' is simply the finite form of the Fubini-Study metric applied to normalized states). There the finite form of the metric is differentiated and then also requires many further lines of derivation to produce the result; here we use the differential form of the metric immediately.

Invariance under unitary transformations follows directly from the definition of a unitary operator as an operator the leaves all inner products of states in $\langle \mathbb{C}^{n+1}, \langle\cdot|\cdot\rangle\rangle$

invariant. The fact that this metric is the unique (upto a constant multiple) metric invariant under all the action of all unitary operators is less simple. The uniqueness argument is based on the fact that $U(n+1)$ acts transitively on $T\mathbb{C}P^n$, for a full explanation see [75].

The FS metric is the projective counterpart to the standard inner product on $\mathbb{C}^{n+1}$; that is, it is compatible with the quotient into rays of $\langle \mathbb{C}^{N+1}, \langle \cdot | \cdot \rangle \rangle$. Every curve $|\psi_t\rangle$ on $\mathbb{C}^{n+1}$ descends to a curve on $\mathbb{C}P^n$. Because of the invariance property of the FS metric, the length any curve in $|\psi_t\rangle$ is unchanged by the transformation $|\psi_t\rangle \mapsto Z_t|\psi_t\rangle$. The tangent vector to $|\psi_t\rangle$, under this transformation, transforms as:

$$|\delta\psi_t\rangle \mapsto \dot{Z}_t|\psi_t\rangle + Z_t|\delta\psi_t\rangle \tag{5.15}$$

The metric transforms as:

$$ds^2 \mapsto \frac{\left(\dot{\bar{Z}}\langle\psi_t| + \bar{Z}\langle\delta\psi_t|\right)\left(\dot{Z}|\psi_t\rangle + Z|\delta\psi_t\rangle\right)}{|Z|^2\langle\psi_t|\psi_t\rangle} \tag{5.16}$$

$$- \frac{\left(\dot{\bar{Z}}\langle\psi_t| + \bar{Z}\langle\delta\psi_t|\right) Z|\psi_t\rangle\langle\psi_t|\bar{Z}\left(\dot{Z}_t|\psi_t\rangle + Z_t|\delta\psi_t\rangle\right)}{\langle\psi_t|\psi_t\rangle^2} \tag{5.17}$$

$$= \frac{|\dot{Z}|^2\langle\psi_t|\psi_t\rangle + \dot{\bar{Z}}Z\langle\psi_t|\delta\psi_t\rangle + \bar{Z}\dot{Z}\langle\delta\psi_t|\psi_t\rangle + |Z|^2\langle\delta\psi_t|\delta\psi_t\rangle}{|Z|^2\langle\psi_t|\psi_t\rangle}$$

$$- \frac{|\dot{Z}|\langle\psi_t|\psi_t\rangle^2 + \dot{\bar{Z}}Z\langle\psi_t|\psi_t\rangle\langle\psi_t|\delta\psi_t\rangle + \bar{Z}\dot{Z}\langle\delta\psi_t|\psi_t\rangle\langle\psi_t|\psi_t\rangle + |Z|^2\langle\delta\psi_t|\psi\rangle\langle\psi|\delta\psi_t\rangle}{|Z|^4\langle\psi_t|\psi_t\rangle^2}$$

$$= ds^2 + \frac{|\dot{Z}|^2\langle\psi_t|\psi_t\rangle + \dot{\bar{Z}}Z\langle\psi_t|\delta\psi_t\rangle + \bar{Z}\dot{Z}\langle\delta\psi_t|\psi_t\rangle}{|Z|^2\langle\psi_t|\psi_t\rangle}$$

$$- \frac{|\dot{Z}|\langle\psi_t|\psi_t\rangle^2 + \dot{\bar{Z}}Z\langle\psi_t|\psi_t\rangle\langle\psi_t|\delta\psi_t\rangle + \bar{Z}\dot{Z}\langle\delta\psi_t|\psi_t\rangle\langle\psi_t|\psi_t\rangle}{|Z|^4\langle\psi_t|\psi_t\rangle^2}$$

$$= ds^2 + 0 = ds^2$$

and is thus unchanged.

A similar approach is possible to prove Fleming's bound [76]:

$$t_\perp \geq \frac{\pi\hbar}{2\Delta\hat{H}} \tag{5.18}$$

using the geodesics of the Killing-Form on $SU(n)$ and an almost identical proof (as the geodesics are the one parameter sub-groups due to bi-invariance).

5.2.3 Proof Techniques For Quantum Speed Limits

This proof of the ML bound is somewhat ad-hoc and I know of no other applications of the inequality applied. While not in and of itself a shortcoming, it would be favourable to seek a general principle by which this statement could be proven and generalised. Furthermore, the proof of the Mandelshtam-Tamm inequality presented above uses a conceptually different approach based on geodesics.

The desired generalisation could be a solution to the question: "Given X amount of a resource Y, what is the least time a quantum system can transition to an orthogonal state". More precisely, by analogy with the two aforementioned bounds, one might endeavour to seek a bound for the form:

$$T_{\perp} \geq \frac{\pi\hbar}{2f(X,Y)} \tag{5.19}$$

where f is a function that remains to be determined. One might also hope for a unifying principle/proof technique for proving such bounds.

5.2.4 More Existing Work On Orthogonality Times

More recent work on bounds on orthogonality times include [64, 68, 69, 77, 74, 78, 79, 80, 81, 50]. Specifically [77, 74, 78] include a role for differential geometry in analysing this aspect of the QSL.

- [79] produces an interesting result generalising the Margolus-Levitin bound to systems to systems with non-unitary dynamics.
- [80] analyses the case of an open driven system and obtains a bound also comparable to the Margolus Levitin bound for non unitary dynamics, a specific model, the damped Jayes-Cummings model is analysed.
- [81] illustrates an application of the Pontryagin minimum principle to the optimal control of $SU(2)$ operators; closed form solutions are obtained as are interesting diagrammatic representations of the optimal trajectories.
- [82] illustrates the absence of a speed limit for quantum systems described by non-Hermitian, PT-symmetric Hamiltonians in a situation where Hermitian quantum mechanics is subject to a finite speed limit.
- [50] discusses the Margolus-Levitin bound in non-Hermitian quantum systems.
- A good discussion of a geometric derivation of the Mandelshtam-Tamm inequality can be found in [77].
- Numerical methods in quantum optimal control are considered in [83] and the Margolus-Levitin bound is shown to be achievable using the Krotov method for deriving control schemes.

5.3 Relativistic And Thermodynamic Limits

Here two relativistic limits to information storage and processing are described. These are included for completeness of context of the physical limits to computation.

5.3.1 Bekenstein Bound

The Bekenstein Bound is a bound on the quantity of information which can be stored in a volume of space, contained within a sphere of radius R, using only a certain amount of energy E. It is briefly reviewed here, for a full discussion see [84, 85, 86]. The bound is given in terms of the entropy of the system within the sphere, this bounds the amount of information storage by limiting the number of distinct states the system within the sphere can possess. The bound is typically stated as:

$$S \leq \frac{2\pi kRE}{\hbar c} \tag{5.20}$$

Here k is Boltzman's constant, $\hbar$ is Planck's constant and c is the speed of light. E is the total mass-energy of the system.

Uncharged, non-rotating Schwarzschild black holes have Bekenstein Hawking entropy which saturate this bound. The entropy of such a black hole is:

$$S_{BH} = \frac{kA}{4} \tag{5.21}$$

where A is the area of the hole's horizon. In this formula A is expressed in terms of the Plank area $\frac{\hbar G}{c^3}$.

5.3.2 Landauer's Principle

Landauer's Principle [87] states that erasing one bit of information always comes with an unavoidable loss of energy of:

$$\delta E = kT \log(2) \tag{5.22}$$

Here E is energy, k is the Boltzman constant and T is the temperature. The original statement, given by Bennet was:

> 'any logically irreversible manipulation of information, such as the erasure of a bit or the merging of two computation paths, must be accompanied by a corresponding entropy increase in non-information bearing degrees of freedom of the information processing apparatus or its environment'

5.4 Comments on Lloyd's 'Ultimate Laptop'

An interesting recent paper [66] sets out to study the following hypothetical object:

> let us calculate the ultimate computational capacity of a computer with a mass of one kilogram occupying a volume of one litre, roughly the size of a conventional laptop. Such a computer, operating at the limits of speed and memory space allowed by physics, will be called the ultimate laptop

Several calculations are made. Lloyd considers the Margolis Levitin bound to represent the minimum time to complete one ‘operation’ within the computer. He further calculates that $E = mc^2 = 8.9874 \times 10^{16}$ joules and then applies the Margolis Levitin bound to determine that 5.4258×10^{50} operations per second are in principle possible in such a machine.

This analysis is however flawed as it is assumed that all the mass energy of the machine is is contributing to the speed of dynamical evolution. This would only be possible for a machine which was itself massless! While no physical principle known to the author explicitly excludes this, it seems highly implausible that such a machine would remain within one litre of space as it operated. Lloyd refers to this problem as “packaging issues”. Also, the Margolis Levitin theorem only limits the time taken to reach an orthogonal state, not to implement specific information processing operations more complex than this elementary one.

Furthermore, the Margolis Levitin theorem is a statement about non-relativistic systems. In this work, the bound is combined in an ad-hoc way with a relativistic calculation of the total energy of the machine. It is not immediately clear that such a calculation is valid, perhaps the relativistic Margolis Levitin bound could be used [88, 89].

The maximum information storage of such a machine is also analysed by appealing to the Bekenstein bound. This is a general relativistic calculation and applies many results from this area. It is estimated that $I = S/kB \ln 2 = 2.13 \times 10^{31}$ bits can be stored in such a machine which results in 10^{19} ops per bit per second. It is again unclear that such a calculation is compatible with the premises of the Margolis Levitin theorem. The system used to store such a number of bits would, as described in the original work, be an uncharged, non-rotating Schwarzschild black holes (in order to saturate the Bekenstein bound). No explanation is offered as to how the processed information would be retrieved from such a system.

Finally, real quantum computers are unlikely to be time independent systems but rather controlled systems.

One step towards producing physically meaningful results about the limitations of computers is to develop methods for calculating the optimal times for implementing specific quantum gates with limited resources. Contributing to this analysis is the main motivation for the work in this thesis.

Part II

Novel Work

Chapter 6

Zermello Navigation On the Special Unitary Group

This chapter discusses the case of Zermelo navigation on $SU(n)$ in order to analyse quantum time optimal control. The problem of driving $\hat{U}_t$, the time evolution operator, from $\hat{U}_0 = \hat{I}$ to a desired gate $\hat{O}$ is considered. In the absence of constraints on the Hamiltonian $\hat{H}$, there would be no lower limit to the time in which this could be achieved. The constraint that $h(i\hat{H}_c(t), i\hat{H}_c(t)) = 1$ for some inner product $h : \mathfrak{su}(n) \times \mathfrak{su}(n) \to \mathbb{R}$, and a generalisation of this, are studied in this chapter.

A solution in the case of a specific class of constraints on the permitted external control fields is given. This work was first presented in [2, 3].

6.1 Zermelo Navigation On $SU(n)$

6.1.1 Right Invariance of h In QM Applications

In order to connect the terminology of quantum dynamics and that of the Zermelo navigation problem we must identify a vector field on $SU(n)$ to play the role of the wind. This can be done by examining the Schrödinger equation 4.1 for a controlled system. The Schrödinger equation for a controlled system 4.6 indicates what the tangent vector to a trajectory of $\hat{U}_t$ is on $SU(n)$, and further that it is comprised of two parts. The first part $-i\hat{H}_0\hat{U}$ represents the system's dynamics in the absence of control, this will play the role of the wind and the associated vector field on $SU(n)$ is $\hat{W}_{\hat{U}} = -i\hat{H}_0\hat{U}$. The second part represents the effect of the control fields, the associated (time dependent) vector field on $SU(n)$ is $-i\hat{H}_c(t)\hat{U}$.

Both these vector fields are right invariant by construction as they are the right translation of some tangent vector at the identity, namely: $-i\hat{H}_0$ and $-i\hat{H}_c$ respectively. Technically, the notion of right translation applied here is not the simplest which acts on group elements, not tangent vectors. The notion applied here is technically the 'canonical lift' of right translation to the tangent bundle $TSU(n)$. That is to say, given the right translation map $R_{\hat{V}} : SU(n) \to SU(n)$ defined on $SU(n)$ by:

$$R_{\hat{V}}\left(\hat{W}\right) = \hat{W}\hat{V} \tag{6.1}$$

one can now define the right translation of any tangent vector $\hat{A}\hat{Y} \in T_{\hat{Y}}SU(n)$ by the map $dR_{\hat{V}}\big|_{\hat{Y}} : T_{\hat{Y}}SU(n) \to T_{\hat{Y}\hat{V}}SU(n)$ which is given by by:

$$dR_{\hat{V}}\big|_{\hat{Y}}\left(\hat{A}\hat{Y}\right) = \hat{A}\hat{W}\hat{V} \tag{6.2}$$

The fact that every tangent vector in $T_{\hat{Y}}SU(n)$ can be expressed as $\hat{A}\hat{Y}$ is a consequence of the fact that all Lie groups are parallelisable (cor. 8.39 [42]). This means that they have 'trivial' tangent bundles. For any Lie group G, TG can be expressed as $TG \cong \mathfrak{g} \times G$ where $\mathfrak{g}$ is the Lie algebra associated to G. In the present case this amounts to $TSU(n) \cong \mathfrak{su}(n) \times SU(n)$ where the $\cong$ symbol refers to a vector bundle isomorphism.

A Riemannian metric h on $SU(n)$ is also needed to set up a navigation problem. This metric represents a limit to the speed of the navigator in the *absence* of wind. Constraints on the control Hamiltonian alone are the only type for which new methods are developed in this thesis. That is to say, some allowed set of control Hamiltonians is prescribed and the control Hamiltonian is restricted to be in this set during any evolution of the controlled system of interest. Mathematically, this is the statement that the constraint does not depend on $\hat{U}_t$ during an evolution, I.e. that the constraint is right invariant.

We now set up the problem of Zermelo navigation on $SU(n)$ and show how it can be applied to quantum mechanics. Suppose that $h : \mathfrak{su}(n) \times \mathfrak{su}(n) \to \mathbb{R}$ is an inner product on $\mathfrak{su}(n)$. Suppose that a controlled quantum system of the form 4.2 is constrained such that its control Hamiltonian $\hat{H}_c(t)$ satisfies $h(i\hat{H}_c(t), i\hat{H}_c(t)) = 1,\ \forall t$. That is, the constraint is time independent and satisfied for all time. It is clear that the right invariance of the Riemannian metric $h_{\hat{U}}$ (formed by right extending h) corresponds to a constraint only on $\hat{H}_t$, not any constraint depending on $\hat{U}_t$ explicitly. This is essentially by the same argumentation for right invariance of a metric representing a constraint in a control problem as that presented in [90].

Suppose further that a drift Hamiltonian $\hat{H}_0$ is given and that $h(i\hat{H}_0, i\hat{H}_0) < 1$. The time evolution operator for our quantum system now satisfies a Schrödinger equation of the form 4.6. The tangent vector $\frac{d\hat{U}_t}{dt}$ to the curve $\hat{U}_t$ has two terms: $-i\hat{H}_c(t)\hat{U}_t$ and $-i\hat{H}_0\hat{U}_t$. In order to fix terminology closer to the original formulation of the Zermelo navigation problem, we define the "wind" vector field on $SU(n)$ by $\hat{W}_{\hat{U}} = -i\hat{H}_0\hat{U}$.

In such a setup, there is enough information to construct the "navigation data" for a Zermelo navigation problem as we now possess a Riemanian metric h according to which the navigator $(\hat{U}_t)$ has speed 1 in the absence of wind and a wind vector field which has h length less less than 1.

6.1.2 The Navigation Randers Metric On $SU(n)$

From these ingredients one can construct a Finsler metric (which is in fact a Randers metric) that has the property that its geodesics are the time optimal trajectories for $\hat{U}_t$ to be driven between given endpoints, by applying Shen's theorem, 2.2.3. This Randers metric is $F_{\hat{U}}(\hat{A}) = \sqrt{\alpha_{\hat{U}}(\hat{A},\hat{A})} + \beta_{\hat{U}}(\hat{A})$. In terms of the navigation data $(h, \hat{W})$ on $SU(n)$, the α and β are found to be:

$$\begin{aligned}\alpha_{\hat{U}}(\hat{A}\hat{U},\hat{A}\hat{U}) &= \frac{\lambda h_{\hat{U}}(\hat{A}\hat{U},\hat{A}\hat{U}) + h_{\hat{U}}(\hat{A}\hat{U},\hat{W}_{\hat{U}})^2}{\lambda^2} \\ &= \frac{h_{\hat{U}}(\hat{A}\hat{U},\hat{A}\hat{U})}{\lambda} + \frac{h_{\hat{U}}(\hat{A}\hat{U},-i\hat{H}_0\hat{U})^2}{\lambda^2} \\ &= \frac{h(\hat{A},\hat{A})}{\lambda} + \frac{h(\hat{A},i\hat{H}_0)^2}{\lambda^2}\end{aligned} \tag{6.3}$$

$$\beta_{\hat{U}}(\hat{A}\hat{U}) = \frac{h_{\hat{U}}(\hat{A}\hat{U},-i\hat{H}_0\hat{U})}{1 - h_{\hat{U}}(-i\hat{H}_0\hat{U},-i\hat{H}_0\hat{U})} = \frac{h(\hat{A},-i\hat{H}_0)}{1 - h(i\hat{H}_0,i\hat{H}_0)} = \frac{h(\hat{A},-i\hat{H}_0)}{\lambda} \tag{6.4}$$

Thus F is, in full, given by:

$$F_{\hat{U}}(\hat{A}\hat{U}) = \sqrt{\frac{h(\hat{A},\hat{A})}{\lambda} + \frac{h(\hat{A},i\hat{H}_0)^2}{\lambda^2}} + \frac{h(\hat{A},i\hat{H}_0)}{\lambda} = F(\hat{A}) \tag{6.5}$$

As stated above, the right invariance of this quantity is clear. Note that λ is a scalar quantity because it is right invariant, and that all right invariant scalar quantities are constant. This is a simplifying factor compared to the case when h and W are right invariant compared to the general case.

6.2 The Case of General h

6.2.1 Euler Poincaré Equations For Optimal Hamiltonians

One could apply the standard Euler-Lagrange (EL) equations to find the geodesics of a Randers metric F on any manifold. The desired geodesics are the stationary curves of the length functional associated to the Lagrangian $\frac{1}{2}F^2$. This would yield, assuming the EL equations for $\frac{1}{2}F^2$ (F^2 rather than F is taken to obtain unit speed geodesics of F) could be solved, the solution for trajectory of $\hat{U}_t$ time optimally connecting $\hat{I}$ to a desired gate $\hat{O}$. The optimal overall Hamiltonian $\hat{H}_t$ could then be found from the Schrödinger equation: $\hat{H}_t = i\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger$. However, another method is now possible due to the right invariance of the metric.

The length functional $L[\hat{U}_t]$ for a curve $\hat{U}_t : [0,T] \to SU(n)$, for a Finlser metric F on $SU(n)$ can be written as follows:

$$L[\hat{U}_t] = \int_{t=0}^{T} F_{\hat{U}_t}\left(\frac{d\hat{U}_t}{dt}\right) dt \tag{6.6}$$

In the case that F is is right invariant one finds:

$$L[\hat{U}_t] = \int_{t=0}^{T} F\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^{-1}\right) dt$$

In the case that $\hat{U}_t$ solves the Schrödinger equation one finds:

$$\begin{aligned} L[\hat{U}_t] &= \int_{t=0}^{T} F_{\hat{U}_t}\left(\frac{d\hat{U}_t}{dt}\right) dt = \int_{t=0}^{T} F_{\hat{U}_t}\left(-i\hat{H}(t)\hat{U}_t\right) \\ &= \int_{t=0}^{T} F\left(-i\hat{H}(t)\right) dt = \int_{t=0}^{T} F\left(-i\hat{H}_0 - i\hat{H}_c(t)\right) dt \end{aligned} \tag{6.7}$$

The length $L[\hat{U}_t]$ depends only on quantities in $\mathfrak{su}(n)$ rather than on the group in general, as all dependence on $\hat{U}_t$ itself has disappeared. In light of this one might expect that it is possible to formulate the geodesic equation for such a Finsler metric as an ODE in $\mathfrak{su}(n)$. This is in fact the case.

In a coordinate-free language (where $\xi \in \mathfrak{su}(n)$) the EP equation reads [91, 92]:

$$\frac{d}{dt}\frac{\partial \ell}{\partial \xi} = -\mathrm{ad}_{\xi}^{*}\left(\frac{\partial \ell}{\partial \xi}\right) \tag{6.8}$$

where $\ell : \mathfrak{su}(n) \to \mathbb{R}$ is the restriction of an arbitrary right invariant Lagrangian $\mathcal{L} : TSU(n) \to \mathbb{R}$ to $\mathfrak{su}(n)$, and ad^* is the co-adjoint representation of $\mathfrak{su}(n)$ [93].

In order to define the adjoint/co-adjoint representation of both $SU(n)$ and $\mathfrak{su}(n)$, first one must introduce a map $\Psi : SU(n) \to \mathrm{Aut}\,(SU(n))$ (wherein $\mathrm{Aut}\,(SU(n))$ means the group of all Lie group automorphisms of $SU(n)$). This map is defined by:

$$\begin{aligned} &\Psi : \hat{U} \mapsto \Psi_{\hat{U}} \\ &\Psi_{\hat{U}} : \hat{V} \to \hat{U}\hat{V}\hat{U}^{-1},\ \forall \hat{V} \in SU(N) \end{aligned} \tag{6.9}$$

and for each $\hat{U}$, $\Psi_{\hat{U}}$ is a Lie group homomorphism. As such, the derivative at the identity $d\left(\Psi_{\hat{U}}\right)\big|_{\hat{I}}$ is a Lie algebra homomorphism.

The adjoint representation of $SU(n)$, $Ad : SU(n) \to \mathrm{Aut}\,(\mathfrak{su}(n))$ (wherein $\mathrm{Aut}\,(\mathfrak{su}(n))$ means the group of all Lie algebra automorphisms of $\mathfrak{su}(n)$) is, as it is for any Lie group, defined by:

$$\begin{aligned} &Ad : \hat{U} \mapsto Ad_{\hat{U}} := d\left(\Psi_{\hat{U}}\right)\big|_{\hat{I}} \\ &Ad_{\hat{U}} : \hat{A} \mapsto \hat{U}\hat{A}\hat{U}^{-1},\ \forall \hat{A} \in \mathfrak{su}(n) \end{aligned} \tag{6.10}$$

The adjoint representation $ad : \mathfrak{su}(n) \to \mathrm{Der}\,(\mathfrak{su}(n))$ of the Lie algebra $\mathfrak{su}(n)$ is the derivatives of all elements in the image Ad at the identity. Here $\mathrm{Der}\,(\mathfrak{su}(n))$ is the Lie algebra of $\mathrm{Aut}\,(SU(n))$. $\mathrm{Der}\,(\mathfrak{su}(n))$ in fact consists of all 'derivations' on $\mathfrak{su}(n)$. To be more explicit a derivation is a map $D : \mathfrak{su}(n) \to \mathfrak{su}(n)$ such that:

$$D\left(\left[\hat{A}, \hat{B}\right]\right) = \left[\hat{A}, D\left(\hat{B}\right)\right] + \left[D\left(\hat{A}\right), \hat{B}\right],\ \forall \hat{A}, \hat{B} \in \mathfrak{su}(n) \tag{6.11}$$

Now one can define ad by:

$$ad = d(Ad)\big|_{\hat{I}} \tag{6.12}$$

This results in the formula:

$$ad_{\hat{A}} : \hat{B} \mapsto \left[\hat{A}, \hat{B}\right], \ \forall \hat{B} \in \mathfrak{su}(n) \tag{6.13}$$

One can readily check that $ad_{\hat{A}}$ is a derivation for any $\hat{A} \in \mathfrak{su}(n)$.

Ad and ad are representations (In both cases the representation space is the Lie algebra $\mathfrak{su}(n)$) of the Lie group $SU(n)$ and Lie algebra $\mathfrak{su}(n)$ respectively as one can readily check that both the following hold:

$$Ad_{\hat{U}} \circ Ad_{\hat{V}} = Ad_{\hat{U}\hat{V}}, \ \forall \hat{U}, \hat{V} \in SU(n) \tag{6.14}$$
$$ad_{[\hat{A},\hat{B}]} = [ad_{\hat{A}}, ad_{\hat{B}}] = ad_{\hat{A}} \circ ad_{\hat{B}} - ad_{\hat{B}} \circ ad_{\hat{A}}, \ \forall \hat{A}, \hat{B} \in \mathfrak{su}(n)$$

The co-adjoint representation of $\mathfrak{su}(n)$ can now be defined in terms of the adjoint representation. The group $SU(n)$ possess a co-adjoint representation also, however this will not be discussed in detail. See [94] for details of the coadjoint representation of a Lie group and Lie Algebra. We use the standard notation for the 'canonical pairing' $\langle \cdot, \cdot \rangle : \mathfrak{su}(n) \times \mathfrak{su}(n)^* \to \mathbb{R}$ ($\mathbb{R}$ as $\mathfrak{su}(n)$ is a real Lie algebra). This pairing is defined by $\langle \hat{A}, f \rangle := f(\hat{A})$. Furthermore, every element of the dual space $\mathfrak{su}(n)^*$ can be represented as $f(\cdot) = \mathrm{Tr}\left(\hat{B}^\dagger \cdot\right)$ for some $\hat{B} \in \mathfrak{su}(n)$. As such, this pairing can be thought of as $\left\langle \hat{A}, Tr\left(\hat{B}^\dagger \cdot\right)\right\rangle = Tr\left(\hat{B}^\dagger \hat{A}\right)$.

The co-adjoint representation ad^* (for which the representation space is $\mathfrak{su}(n)^*$) is defined implicitly via:

$$\left\langle \hat{C}, ad_{\hat{A}}\left(\hat{B}\right)\right\rangle = -\left\langle ad^*_{\hat{A}}\left(\hat{C}\right), \hat{B}\right\rangle \tag{6.15}$$

For clarity, each $ad^*_{\hat{A}} : \mathfrak{su}(n)^* \to \mathfrak{su}(n)^*$, for each $\hat{A} \in \mathfrak{su}(n)$. Note the minus sign in 6.8: this is due the the metric being right invariant rather than left as is more commonly studied in pure mathematics contexts. There are some additional conditions on $\mathcal{L}$ for the EP equations to apply; these can be readily found in any mathematical description of the theory of Lagrangian reduction [95, 91]. It is clear that all Finsler metrics meet the required conditions. For example, it is clear that the regularity condition is met, as it is present in the definition of a Finsler metric.

This equation may also been seen with a δ (signifying a functional derivative) in place of the d above; this is the form of the equation which applies to infinite dimensional problems rather than the finite dimensional ones studied here.

On fixing a basis $\{\hat{B}_k\}$ for $\mathfrak{su}(n)$ and expressing an arbitrary element $-i\hat{H}_t$ as $\xi^k \hat{B}_k$, the EP equation takes the form [91, 92]:

$$\frac{d}{dt}\frac{\partial \ell}{\partial \xi^d} = -C^b_{ad}\frac{\partial \ell}{\partial \xi^b}\xi^a \tag{6.16}$$

where C^d_{ab} are the structure constants of $\mathfrak{su}(n)$. See [39] for details of structure constants in general, and [96] for $\mathfrak{su}(n)$ specifically, where the structure constants of $\mathfrak{su}(2)$ and $\mathfrak{su}(4)$ are given explicitly. The tensor C possesses many symmetries, including $C^a_{bd} = -C^a_{db}$ for example; this follows directly from the antisymmetry of the Lie bracket. As of yet, the author has not found a way to exploit these symmetries as a tool for simplifying the EP equations in the case of $\mathfrak{su}(n)$.

Henceforth the subscripts indicating a point on $SU(n)$ are dropped from α and β, and they are understood to be restricted to the tangent space of $SU(n)$ at the identity, i.e. $\mathfrak{su}(n)$. However, coordinate indicies still appear.

This procedure can be applied to finding the geodesics of a right invariant Randers metrics on $SU(n)$. The variable ξ, in the present case, takes the form $\xi = \frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger = -i\hat{H}_t$. Thus, the EP equation for a Randers metric $\frac{1}{2}F^2$ on $SU(n)$ is satisfied by the overall Hamiltonian driving $\hat{U}_t$ along a geodesic of F, i.e. a time optimal trajectory.

Setting ℓ to be the square, to obtain unit speed geodesics, of a Randers norm $\ell(\hat{A}) = \frac{1}{2}(F\big|_{\hat{I}}(\hat{A}))^2 = \frac{1}{2}\left(\sqrt{\alpha(\hat{A},\hat{A})} + \beta(\hat{A})\right)^2$, i.e. the restriction of a Randers metric F on $SU(n)$ to $\mathfrak{su}(n)$, we can derive the EP equation associated to the geodesics of F. Substituting into the EP equation one finds:

$$\frac{\partial \ell}{\partial \xi^d} = \frac{1}{2}\left((\alpha_{ij}\xi^i\xi^j)^{1/2} + \beta_k\xi^k\right)\left(||\xi||_\alpha^{-1}\alpha_{nd}\xi^n + \beta_d\right) \tag{6.17}$$

then differentiating one finds:

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial \ell}{\partial \xi^d}\right) &= \frac{d}{dt}\left(||\xi||_\alpha + \beta_k\xi^k\right)\left(||\xi||_\alpha^{-1}\alpha_{nd}\xi^n + \beta_d\right) \\ &= \left(||\xi||_\alpha^{-1}\langle\dot{\xi},\xi\rangle_\alpha + \beta_j\dot{\xi}^j\right)\left(||\xi||_\alpha^{-1}\alpha_{md}\xi^m + \beta_d\right) - \\ &\quad \left(||\xi||_\alpha + \beta_k\xi^k\right)\left(||\xi||_\alpha^{-3}\langle\xi,\dot{\xi}\rangle_\alpha\alpha_{hd}\xi^h - ||\xi||_\alpha^{-1}\alpha_{kd}\dot{\xi}^k\right)\end{aligned} \tag{6.18}$$

These yield the EP equation of a geodesic:

$$\begin{aligned}&\left(||\xi||_\alpha^{-1}\langle\dot{\xi},\xi\rangle_\alpha + \beta_j\dot{\xi}^j\right)\left(||\xi||_\alpha^{-1}\alpha_{md}\xi^m + \beta_d\right) - \\ &\left(||\xi||_\alpha + \beta_k\xi^k\right)\left(||\xi||_\alpha^{-3}\langle\xi,\dot{\xi}\rangle_\alpha\alpha_{hd}\xi^h - ||\xi||_\alpha^{-1}\alpha_{kd}\dot{\xi}^k\right) = \\ &\qquad -C^a_{bd}\left(\left(||\xi||_\alpha + \beta_k\xi^k\right)\left(||\xi||_\alpha^{-1}\alpha_{na}\xi^n + \beta_a\right)\right)\xi^b\end{aligned} \tag{6.19}$$

where we take the following meanings: $||\xi||_\alpha := \sqrt{\alpha_{ij}\xi^i\xi^j}$ and $\langle\mu,\nu\rangle_\alpha := \alpha_{ij}\mu^i\nu^j$.

We are interested in the geodesics associated to a navigation problem specified in terms of its navigation data. Such an equation can, in principle, be obtained by substituting in the definitions of α and β in terms of h and W from Shen's solution to the navigation problem described in 2.3. As of yet the author has not been able to obtain a tractable form for this equation.

6.3 Using Robles Theorem To Determine Geodesics

6.3.1 The Navigation Metric In the Bi-Invariant Case

The case studied in this section is that in which h is bi-invariant, i.e. $h(i\hat{A}, i\hat{B}) = \kappa \operatorname{Tr}(\hat{A}\hat{B})$. This is the unique bi-invariant metric (up to a choice of positive $\kappa \in \mathbb{R}$) defined by the right (or left) translation of the Killing form from the identity. This case was treated in the work [2]. For a clear review of Killing forms and bi-invariant metric on Lie groups see ([97], §14.2).

The navigation Randers metric F 6.5 evaluates, after some elementary algebra, to:

$$F(i\hat{A}\hat{U}) = \frac{1}{\rho - 1} \frac{\operatorname{Tr}(\hat{A}\hat{H}_0)}{\operatorname{Tr}(\hat{H}_0^2)} \left(1 \pm \sqrt{1 + (\rho - 1) \frac{\operatorname{Tr}(\hat{H}_0^2) \operatorname{Tr}(\hat{A}^2)}{(\operatorname{Tr}(\hat{A}\hat{H}_0))^2}} \right) \tag{6.20}$$

where $i\hat{A}\hat{U} \in T_{\hat{U}}SU(N)$ and:

$$\rho := \frac{\operatorname{Tr}(\hat{H}_c^2(t))}{\operatorname{Tr}(\hat{H}_0^2)} > 1 \tag{6.21}$$

This metric is obtained by substituting $h(i\hat{A}, i\hat{B}) = \kappa \operatorname{Tr}\left(\hat{A}\hat{B}\right)$ into 6.5. κ does not appear directly in F as it has been absorbed into the value of ρ.

6.3.2 Geodesics and Optimal Hamiltonians In the Bi-Invariant Case

Here I note that the centeral result of this section also obtained my Brody and Meier [98] concurrently with this work by a different method.

The geodesics of 6.20 can be determined by an application of a special case of Robles theorem [99, thm.2]. We use $\sigma = 0$ in that theorem, as the special case of a Killing field (see [100] for definitions) in place of the infinitesimal homothety. The definition of a Killing field is given here for clarity:

Definition 6.3.1. A smooth vector field $X \in \Gamma(TM)$ on a Riemanian manifold (M, g) is Killing field of g iff:

$$\pounds_X g = 0 \tag{6.22}$$

where $\pounds$ is the Lie derivative. This is equivalent, in terms of the Levi-Civita connection of g, to:

$$g(\nabla_x, Z) + g(Y, \nabla_X Z) = 0 \tag{6.23}$$

for all vectors $X, Y, Z \in T_pM$ for each $p \in M$. In local coordinates this condition is equivalent to Killing's equation:

$$\nabla_\mu X_\nu + \nabla_\nu X_\mu = 0 \tag{6.24}$$

We specialise to $SU(n)$, rather than a general manifold, as this is the case relevant to quantum mechanics. In fact, there are no infinitesimal homotheties that are not Killing fields for the bi-invariant metric on $SU(n)$, so no real restriction has been incurred on which metrics can have their geodesics determined using the following theorem. The theorem states:

Theorem 6.3.1 (adapted from [99, thm.2]). *Given:*

- *A Riemannian manifold* (M, h)
- *A smooth vector field* $\hat{W} \in \Gamma(TSU(n))$ *on* $SU(n)$ *such that* $\pounds_{\hat{W}}(h) = 0$ *(that is, the Lie derivative of the metric is* 0*, or equivalently* $\hat{W}$ *is a Killing vector field).*

Given that F *is the Randers metric solving the Zermelo navigation problem on* M *for navigation data* h *and* $\hat{W}$*, then the unit* F *speed geodesics of* F *are given by* $\hat{V}_t = \phi_t(\hat{S}_t)$*, where:*

- ϕ_t *is the flow associated to* $\hat{W}$
- $\hat{S}$ *is a unit speed geodesic of* h

Furthermore, any geodesic of F *obtained this way is a global length minimiser if and only the associated Riemannian geodesic of* h *is a length minimiser of* h *[99].*

In the case that h is the bi-invariant metric, the unit speed geodesics $\hat{S}_t$ are the one parameter subgroups of $SU(n)$, parametrised to have unit h speed. These can all be expressed as $\hat{S}_t = \exp(it\hat{D})$ for some $\hat{D} \in \mathfrak{su}(n)$ that is a unit vector for the same h. The flow on $SU(n)$ associated to the vector field $\hat{W}_{\hat{U}} = -i\hat{H}_0\hat{U}$ is:

$$\phi_t(\hat{U}) = \exp(-it\hat{H}_0)\hat{U} \tag{6.25}$$

This follows from the observation that the equation defining the flow:

$$\frac{d\phi_t(\hat{U})}{dt} = -i\hat{H}_0\phi_t(\hat{U}) \tag{6.26}$$

is exactly the Schrödinger equation with Hamiltonian $\hat{H}_0$. We thus conclude that the time optimal trajectories are given by:

$$\begin{aligned} \hat{U}_t = \phi_t(\hat{S}_t) &= \phi_t\left(\exp(it\hat{D})\right) \\ &= \exp(-it\hat{H}_0)\exp(it\hat{D}) \end{aligned} \tag{6.27}$$

This is to be compared with [4, eqn.51] which exhibits a similar product of exponentials structure.

We determine the optimal Hamiltonian by assuming $\hat{U}_t$, a geodesic, solves the Schrödinger equation for an as yet unknown Hamiltonian $\hat{H}_t$:

$$\frac{d\hat{U}_t}{dt} = -i\hat{H}_t\hat{U}_t \tag{6.28}$$

which implies that:

$$
\begin{aligned}
\hat{H}_t &= i\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger \\
&= i\left((-i\hat{H}_0)\hat{V}_t + \hat{V}_t(i\hat{D})\right)\hat{V}^\dagger \\
&= i\left((-i\hat{H}_0) + \hat{V}_t(i\hat{D})\hat{V}^\dagger\right) \\
&= \hat{H}_0 - \hat{V}_t(\hat{D})\hat{V}^\dagger \\
&= \hat{H}_0 - \exp(-it\hat{H}_0)(\hat{D})\exp(it\hat{H}_0) \\
&= \hat{H}_0 + i\mathrm{Ad}_{\exp(-it\hat{H}_0)}(i\hat{D})
\end{aligned}
\tag{6.29}
$$

Finally, in order to conclude that the curves in 6.27 are the geodesics and that their associated Hamiltonians are given by 6.29, we check that the given $\hat{W}_{\hat{U}}$ is a Killing field for the metric h. This is achieved by checking $\pounds_{\hat{W}}(h) = 0$ thus:

$$
\begin{aligned}
&\frac{d}{dt}h_{\phi_t(\hat{V})}\left(d\phi_t\big|_{\hat{V}}(\hat{A}\hat{V}), d\phi_t\big|_{\hat{V}}(\hat{A}\hat{V})\right)\Bigg|_{t=0} \\
&= \frac{d}{dt}h_{\exp(-it\hat{H}_0)\hat{V}}\left(\exp(-it\hat{H}_0)\hat{A}\hat{V}, \exp(-it\hat{H}_0)\hat{A}\hat{V}\right)\Bigg|_{t=0} \\
&= \frac{d}{dt}h_{\exp(-it\hat{H}_0)}\left(\exp(-it\hat{H}_0)\hat{A}, \exp(-it\hat{H}_0)\hat{A}\right)\Bigg|_{t=0} \\
&= \frac{d}{dt}h_{\hat{I}}(\hat{A},\hat{A})\Bigg|_{t=0} = 0
\end{aligned}
\tag{6.30}
$$

where $\hat{V}$ is an arbitrary group element and $\hat{A}\hat{V}$ is an arbitrary element of $T_{\hat{V}}SU(n)$. We have $d\phi_t\big|_{\hat{V}}(\hat{A}\hat{V}) = \exp(-it\hat{H}_0)\hat{A}\hat{V}$ trivially, as it is the differential of a linear map. Here, both the left and the right invariance of the metric h have been appealed to; this proof would need to be modified, or simply does not hold, in the case that h is not the unique bi-invariant metric. In such a scenario one would have to solve 6.19 directly to obtain the optimal Hamiltonian.

6.3.3 Optimal Gate Times From Geodesic Lengths

What remains to determine is the formula for a geodesic with desired endpoints (connecting the identity $\hat{I}$ to a desired operator $\hat{O} \in SU(n)$) and the corresponding Hamiltonian driving $\hat{U}_t$ along the geodesic. This boils down to determining the $\hat{D}$ corresponding to a given $\hat{O} \in SU(n)$. In the bi-invariant case, to determine which $\hat{D}$ yields the geodesic with endpoints $\hat{I}$ and $\hat{O}$ such that the system traverses the geodesic in time T, we need to solve:

$$
\hat{U}_T = \exp(-iT\hat{H}_0)\exp(iT\hat{D}) = \hat{O} \tag{6.31}
$$

Rearranging and taking matrix logs:

$$
\begin{aligned}
\exp(iT\hat{D}) &= \exp(iT\hat{H}_0)\hat{O} \\
i\hat{D} &= \frac{1}{T}\log\left(\exp(iT\hat{H}_0)\hat{O}\right)
\end{aligned}
\tag{6.32}
$$

which yields the desired geodesic and corresponding control Hamiltonian:

$$\hat{U}_t = \exp(-it\hat{H}_0)\exp\left(\frac{t}{T}\log\left(\exp(iT\hat{H}_0)\hat{O}\right)\right) \tag{6.33}$$
$$= \exp(-it\hat{H}_0)\left(\exp(iT\hat{H}_0)\hat{O}\right)^{t/T}$$

$$\hat{H}_c(t) = \frac{i}{T}\exp(-it\hat{H}_0)\log\left(\exp(iT\hat{H}_0)\hat{O}\right)\exp(it\hat{H}_0) \tag{6.34}$$
$$= \frac{i}{T}\log\left(\exp(-it\hat{H}_0)\exp(iT\hat{H}_0)\hat{O}\exp(it\hat{H}_0)\right)$$
$$= \frac{i}{T}\log\left(\exp\left(i(T-t)\hat{H}_0\right)\hat{O}\exp(it\hat{H}_0)\right)$$

We can take the $\exp(\pm it\hat{H}_0)$ factors inside the logarithm, because the matrix logarithm is analytic [101, Ch.7], which follows from the fact that any matrix function f which is defined by a power series obeys $f(\hat{V}^{-1}\hat{A}\hat{V}) = \hat{V}^{-1}f(\hat{A})\hat{V}$ for all matrices $\hat{A}$ and all non singular $\hat{V}$.

In order to determine the optimal time T_{opt} to implement a gate $\hat{O}$ in systems constrained such that $h(i\hat{H}_c(t), i\hat{H}_c(t)) = 1$ one must find the F (solving the navigation problem) length of the F geodesic connecting $\hat{I}$ to $\hat{O}$. Insisting that the left hand side of 6.32 has norm 1 according to h, that $h(i\hat{D}, i\hat{D}) = 1$, in accordance with the premise that it is the *unit speed* geodesics of h that are needed, we determine that:

$$1 = h(i\hat{D}, i\hat{D}) = h\left(\frac{1}{T}\log\left(\exp(iT\hat{H}_0)\hat{O}\right), \frac{1}{T}\log\left(\exp(iT\hat{H}_0)\hat{O}\right)\right) \tag{6.35}$$

which yields the following equation to be solved for T:

$$-\frac{\kappa}{T^2}\operatorname{Tr}\left(\log\left(\exp(iT\hat{H}_0)\hat{O}\right)^2\right) = 1 \tag{6.36}$$

The smallest positive solution is the optimal time; we refer to this as T_{opt}. At the time of writing, we have found no method for solving this analytically in general; it appears prohibitively difficult by standard means known to the authors. However, once $\hat{H}_0$ and $\hat{O}$ are given, it can easily be solved numerically; some simple cases are illustrated in 7.

Once T_{opt} is known, either analytically or numerically, then the true geodesics and corresponding optimal control Hamiltonian are:

$$\hat{U}_t = \exp(-it\hat{H}_0)\left(\exp(iT_{\text{opt}}\hat{H}_0)\hat{O}\right)^{t/T_{\text{opt}}} \tag{6.37}$$

$$\hat{H}_c(t) = \frac{i}{T_{\text{opt}}}\log\left(\exp\left(i(T_{\text{opt}}-t)\hat{H}_0\right)\hat{O}\exp(it\hat{H}_0)\right)$$

Obtaining an Approximate Formula For $\hat{D}$

We can use the well-known BCH formula [39, §3] to evaluate approximations to $i\hat{D}$ as it provides a series type representation for the solution to $\exp(z) = \exp(x)\exp(y)$. Given a certain x, y in a Lie algebra the solution for z is given by:

$$z = x + y + \frac{1}{2}[x,y] + \frac{1}{12}[x,[x,y]] - \frac{1}{12}[y,[x,y]] - \frac{1}{24}[y,[x,[x,y]]] + \cdots \tag{6.38}$$

We apply this to 6.32 to solve $\exp(iT\hat{D}) = \exp(iT\hat{H}_0)\exp(\log(\hat{O}))$, to obtain:

$$i\hat{D} = i\hat{H}_0 + \frac{1}{T}\log\hat{O} + \frac{1}{2}[i\hat{H}_0, \log(\hat{O})] + \frac{T}{12}[i\hat{H}_0, [i\hat{H}_0, \log(\hat{O})]] \\ - \frac{1}{12}[\log\hat{O}, [i\hat{H}_0, \log\hat{O}]] - \frac{T}{24}[\log(\hat{O}), [i\hat{H}_0, [i\hat{H}_0, \log(\hat{O})]]] + \cdots \tag{6.39}$$

A Simplifying Case That Can Be Solved In Closed Form

One special case that can be solved analytically is where $\hat{O}$ and $\hat{H}_0$ commute. Expanding the matrix logarithm using $\log(\hat{A}\hat{B}) = \log(\hat{A}) + \log(\hat{B})$, rearranging and applying the standard quadratic formula gives:

$$T_{\text{opt}} = \frac{i\kappa\,\mathrm{Tr}(\hat{H}_0\log(\hat{O}))}{k\,\mathrm{Tr}(\hat{H}_0^2) - 1} \pm \frac{1}{2}\sqrt{\frac{\kappa\,\mathrm{Tr}((\log(\hat{O}))^2)}{\kappa\,\mathrm{Tr}(\hat{H}_0^2) - 1} - \frac{\kappa^2\,\mathrm{Tr}(\hat{H}_0\log(\hat{O}))^2}{(\kappa\,\mathrm{Tr}(\hat{H}_0^2) - 1)^2}} \tag{6.40}$$

where, as in [2], the $\pm$ is chosen to ensure a positive time.

6.4 Generalising To a Non-Riemanian Bi-Invariant Finsler Metric $\check{F}$ In Place of h

A generalisation of Robel's theorem holds for general bi-invariant Finsler Metric $\check{F}$ [100]. In this section this generalisation is applied to generalise the results of previous sections 6.33.

It is worth noting what is *not* achieved by studying Zermelo navigation on only Riemannian manifolds rather than more general Finsler manifolds. The only constraints on quantum systems that can be studied this way are those which restrict the control Hamiltonian $i\hat{H}_c$ to be restricted to the unit sphere of a norm induced by a given inner product h on $\mathfrak{su}(n)$. This is because such inner-products are in one-to-one correspondence with right invariant metrics on $SU(n)$ (by right translation). This allows only quadratic constraints to be studied.

However, not all physically meaningful constants are of this form. An interesting constraint not included in this class is the restriction that the energy expectation (in some specific state $|\psi\rangle$) associated to the control Hamiltonian alone is equal to a fixed constant for all time:

$$\kappa\langle\psi|\hat{H}_c(t) - \frac{1}{D}E_0(t)\hat{I}|\psi\rangle = 1 \tag{6.41}$$

Choosing $|\psi\rangle$ to be the uniform superposition state (as described in [2]) this constraint evaluates to:

$$\kappa\,\mathrm{Tr}\left(\hat{H}_c(t) - \frac{1}{D}E_0(t)\hat{I}\right) = -\kappa E_0(t) = 1 \tag{6.42}$$

where $E_0(t)$ is the lowest eigenvalue of $\hat{H}_c(t)$ and D is the dimension of the vector space on which $\hat{H}_c$ acts.

The main result of [100] is particularly relevant as it generalises [99] and allows one to replace the role of h with a general Finsler metric that has $i\hat{H}_0\hat{U}$ as a Killing field. One class of Finsler metrics with this property is the bi-invariant ones, of which there are many. The proof of bi-invariance is similar to the bi-invariant Riemannian cases already presented, and so is omitted. Examples of such constraints are found in the Finsler metrics formed from the right translation of the Shatten p-Norms on $\mathfrak{su}(n)$. These correspond the constraint that:

$$F^{(p)}(\hat{H}_c(t)) := \kappa \left(\sum_n |E_n|^p \right)^{1/p} = 1 \quad \forall t \tag{6.43}$$

thus generalising the case of the bi-invariant Riemannian metric studied above, wherein $p = 2$ (the only value of p yielding a Riemannian metric on $SU(n)$). Solving the navigation problem in general, to obtain a closed form for the navigation metric, has not been achieved by the mathematics community, as far as the authors are aware. However, there are other cases besides the Riemannian case that have been solved; the Kropina metric case [102] is notable. In the absence of a solution to the navigation problem analogous to the role of Randers metrics in the Riemannian case, alternative methods must been sought. The central result of [100] allows one to determine the geodesics of the Finsler metric solving the navigation problem on $SU(n)$ for which $\hat{H}_c(t)$ is constrained such that $F^{(p)}(\hat{H}_c(t)) = 1$. The geodesics of the (unknown) metric solving the navigation problem in such a generalised case are:

$$\hat{V}_t = \exp(-it\hat{H}_0)\exp(it\hat{D}) \tag{6.44}$$

where $F^{(p)}(i\hat{D}) = 1$.

One can also find the time optimal Hamiltonian for implementing gate $\hat{O}$ to be the same as the Riemannian ($p = 2$) case:

$$\hat{H}_c(t) = \frac{i}{T}\exp(-it\hat{H}_0)\log\left(\exp(iT\hat{H}_0)\hat{O}\right)\exp(it\hat{H}_0) \tag{6.45}$$

except with T_{opt} taking a different value. The requirement is now that T_{opt} is the smallest value of T that solves:

$$F^{(p)}\left(\log(\exp(iT\hat{H}_0)\hat{O})\right) = T \tag{6.46}$$

As in the Riemannian case, the authors do not know of a method for solving this in closed form and a numerical method mus be used.

6.5 Advantages of Method

The method in [4, 5] can produce the optimal trajectory and even the optimal (over all trajectories) time to implement a gate. However, a primary advantage (over the work of Carlini et. al [4, 5]) of the method presented in this chapter is that one can determine the optimal time for $\hat{U}_t$ to traverse an arbitrary curve on $SU(n)$. Previous methods cannot achieve this. This can be achieved as the optimal time for a system (meeting the required premises) to traverse a given curve on $SU(n)$ is the length of that curve according to the metric solving the navigation problem as described above.

6.5.1 Time Independent Trajectories

As an example of finding the time required to implement a gate using trajectories only of a prescribed type can be restricting to constant control fields. In such a situation a closed form, explicit, formula (unlike eqn. 6.36 which is implicit) for the time required to implement an arbitrary gate can be found.

In order to determine the time required to implement a gate $\hat{O} \in SU(n)$ in such a system one must first find the curve connecting the identity $\hat{I}$ to $\hat{O}$. The assumption that the controls are constant results in the the overall Hamiltonian being constant:

$$\hat{H} = \hat{H}_0 + \hat{H}_c \tag{6.47}$$

where $\hat{H}_c$ is a constant. In such a scenario the system is described by the standard Schrödinger equation with time independent Hamiltonian:

$$\frac{d\hat{U}_t}{dt} = -i\hat{H}\hat{U}_t \tag{6.48}$$

This equation is solved by a matrix exponential:

$$\hat{U}_t = e^{-it\hat{H}} \tag{6.49}$$

as can be readily confirmed. Another method for obtaining the form of the time evolution operator is to apply Stone's theorem [103]. These trajectories are the one-parameter subgroups of $SU(n)$.

One can find the Hamiltonian driving $\hat{U}_t$ along a one-parameter subgroup and reaching $\hat{O}$ in time T by setting:

$$e^{-iT\hat{H}} = \hat{O} \tag{6.50}$$

and the taking the matrix log of both sides, for a full discussion of the definition and ambiguity of the matrix logarithm see 6.5.1. Here, the principal logarithm is chosen. This results in:

$$\log\left(e^{-iT\hat{H}}\right) = \log\left(\hat{O}\right) \tag{6.51}$$

From which one can conclude:

$$-iT\hat{H} = \log\left(\hat{O}\right) \tag{6.52}$$

and thus:

$$\hat{H} = \frac{i}{T}\log\left(\hat{O}\right) \tag{6.53}$$

Before the result 6.5 of this chapter can be applied to such a system, a constraint on the 'size' of $\hat{H}_c$ must also be imposed. We assume that $h\left(i\hat{H}_c, i\hat{H}_c\right) = 1$ for some inner product h on $\mathfrak{su}(n)$. We further assume that $h\left(i\hat{h}_0, i\hat{H}_0\right) < 1$. These assumptions assure that the premises of 6.5 are met and that, subsequently, the desired time is the F length of the curve:

$$\hat{U}_t = e^{\frac{t}{T}\log(\hat{O})} \tag{6.54}$$

according to the right invariant Randers metric:

$$F_{\hat{V}}(\hat{A}\hat{V}) = \sqrt{\frac{h(\hat{A},\hat{A})}{1-h(i\hat{H}_0, i\hat{H}_0)} + \frac{\left(h(\hat{A}, i\hat{H}_0)\right)^2}{\left(1-h(i\hat{H}_0, i\hat{H}_0)\right)^2}} + \frac{h(\hat{A}, i\hat{H}_0)}{1-h(i\hat{H}_0, i\hat{H}_0)} \tag{6.55}$$

The required quantity (as F is right invariant) is:

$$\begin{aligned} T_{opt} = L[\hat{U}_t] &= \int_0^T F\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger\right) dt \\ &= \int_0^T F\left(\frac{1}{T}\log\left(\hat{O}\right)\right) dt \\ &= F\left(\log\left(\hat{O}\right)\right) \end{aligned} \tag{6.56}$$

Ambiguity of the Matrix Logarithm

For a clear discussion of this material in a more general context see [104]. The matrix logarithm, unlike the matrix exponential, is not in general unique. The log of a square complex matrix $\hat{A}$ is defined as the multivalued inverse of the matrix exponential. This is to say: any matrix $\hat{B}$ such that:

$$\exp(\hat{B}) = \hat{A} \tag{6.57}$$

is called *a* logarithm of $\hat{A}$ and one has $\exp(\log(\hat{A})) = \hat{A}$. Within the general linear group $GL(n, \mathbb{C})$ ($GL(n)$ henceforth where no confusion is possible) of non-singular complex matrices the issue of existence of a logarithm is particularly simple. All non-singular complex matrices have at least one logarithm in $\mathfrak{gl}(n)$, the Lie algebra of $GL(n)$ which consists of all complex square n by n square matrices. Note that, in contrast to the log of positive real numbers, this is not equivalent to $\log(\exp(\hat{A}))) = \hat{A}$ which does not hold for all logarithms of $\hat{A}$.

As unitary matrices are all invertible, they each have at least a single logarithm. One of these logarithms is singled out, the *principal logarithm*. This is the unique logarithm such that it's eigenvalues all lie on the strip in the complex plane:

$$\{Z \in \mathbb{C} \;\text{ s.t.} -\pi < \Im(Z) < \pi\} \tag{6.58}$$

Where no specific matrix logarithm is mentioned, in this thesis the principal log is intended. In the case of a unitary matrix, the situation is particularly simple [104] as the logarithm will always have purely imaginary eigenvalues as it is skew Hermitian. In every optimal time examined in this thesis in which a matrix logarithm appears, the principle log appears to give to least time. However, firmly establishing this will require further work.

Chapter 7

Example Calculations Of Optimal Times

In this chapter, example calculations of optimal times for implimenting quantum gates are given. This work was first published in [3].

7.1 Example: A Spin In a Magnetic Field With Constant Control Fields

In order to exemplify the concept of finding the time to traverse a specific trajectory the example of a single spin (i.e. a spin $\frac{1}{2}$ particle or nucleus) is studied.

The drift Hamiltonian is taken to be:

$$\hat{H}_0 = B^x\sigma^x + B^y\sigma^y \tag{7.1}$$

The physical interpretation of this is that the field $\vec{B} = \begin{pmatrix} B_x \\ B_y \\ 0 \end{pmatrix}$ represents the effect of a constant (in time) magnetic field outside the control of the experimenter. The notation $B^2 = \vec{B} \cdot \vec{B} = (B^x)^2 + (B^y)^2$ will be employed as will similar notation for other vectors.

We further assume that there is another magnetic field under the control of the experimenter, $D = \begin{pmatrix} D_x \\ D_y \\ D_z \end{pmatrix}$. The control Hamiltonian is given by: $\hat{H}_c = D^x\sigma^x + D^y\sigma^y + D^z\sigma^z$. We do, however, assume that the experimenter can only implement fields of limited 'strength'. The notion of strength that we choose here is a very physically reasonable one, we choose to restrict $\kappa\,\mathrm{Tr}(\hat{H}_c^2) = 1$ for some $\kappa \in \mathbb{R}^+$.

The requirement that $\kappa\,\mathrm{Tr}(\hat{H}_c^2) = 1$ can be evaluated by applying the Clifford algebra (of $\mathbb{R}^3$ with the standard euclidean metric) property of the Pauli matrices σ^k [105]. This is: $(D_k\sigma^k)^2 = (\vec{D} \cdot \vec{D})\hat{I}$. This implies:

$$\begin{aligned} \mathrm{Tr}(H_c^2) &= \mathrm{Tr}\left((D_k\sigma^k)^2\right) \\ &= \mathrm{Tr}((\vec{D} \cdot \vec{D})\hat{I}) = 2\vec{D} \cdot \vec{D} \\ &= 2(D_x^2 + D_y^2 + D_z^2) = \frac{1}{\kappa} \end{aligned} \tag{7.2}$$

From this we conclude that, in terms of the physical quantities directly, the constraint reads: $D^2 = 1/2\kappa$.

Similarly, the requirement that $\kappa \,\mathrm{Tr}(\hat{H_0}^2) < 1$ can be evaluated:

$$\mathrm{Tr}(H_0^2) = \mathrm{Tr}((B_x\sigma^x + B_y\sigma^y)^2) = 2B^2 < \frac{1}{\kappa} \tag{7.3}$$

where $B^2 := B_x^2 + B_y^2$. Equations 7.2 and 7.3 give $B^2 < D^2$; the control field overcomes the drift field.

In the specific case of h one can now conclude, after some simple algebraic manipulation of 6.55 and 6.56, that:

$$T_{\mathrm{opt}} = \frac{1}{\rho - 1}\frac{i\,\mathrm{Tr}(\hat{H}_0 \log(\hat{O}))}{\mathrm{Tr}(\hat{H}_0^2)}\left(1 \pm \sqrt{1 + (\rho-1)\frac{\mathrm{Tr}(\hat{H}_0^2)\,\mathrm{Tr}((\log(\hat{O}))^2)}{\left(\mathrm{Tr}(\hat{H}_0 \log(\hat{O}))\right)^2}}\right) \tag{7.4}$$

$\mathrm{Tr}(\hat{H}_0 \log(\hat{O}))$ is always purely imaginary, and thus the expression evaluates to a real result, despite the presence of i. The choice of $\pm$ is made to ensure positivity. Within the formula 7.4 we have defined:

$$\rho := \frac{\mathrm{Tr}(\hat{H}_c^2(t))}{\mathrm{Tr}(\hat{H}_0^2)} > 1 \tag{7.5}$$

for convenience.

As an example, some particular operation $\hat{O}$ is chosen. We can then calculate its optimal implementation time. Setting $\hat{O} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = -i\sigma^y$ gives a gate that sends each of the two computational basis states to an orthogonal state. We then find the optimal implementation time thus:

1. $\rho = D^2/B^2$
2. $\log(\hat{O}) = \log \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{\pi}{2} \\ \frac{\pi}{2} & 0 \end{pmatrix} = -i\frac{\pi}{2}\sigma^y$
3. $\mathrm{Tr}\left(\log(\hat{O})\hat{H}_0\right) = -\pi i B_y$
4. $\mathrm{Tr}\left((\log(\hat{O}))^2\right) = -\pi^2/2$

Substituting into 7.4, one can find the time for this gate: Combining these terms and substituting into 7.4 yields:

$$T_{\mathrm{opt}} = \frac{\pi}{2}\frac{B_y}{(D^2 - B^2)}\left(1 \pm \sqrt{1 + \frac{D^2 - B^2}{B_y^2}}\right) \tag{7.6}$$

7.2 A Two Level System: Constant vs Non-Constant Controls

In order to exemplify a direct comparison between the speed limits for two different types of trajectory, a specific example of the same system 7.1 is considered but with non-constant controls permitted.

The 'size' constraint on the control Hamiltonian is again $\kappa \operatorname{Tr}\left(\hat{H}_c(t)^2\right) = 1$. However, now this constraint is imposed for all times and the control Hamiltonian $\kappa \hat{H}_c(t)$ is permitted to be time dependent. The drift Hamiltonian is as before: $\hat{H}_0 = B^x\sigma_x + B^y\sigma_y$ and is further such that $\kappa \operatorname{Tr}\left(\hat{H}_0^2\right) \leq 1$. By applying formula 6.36 one finds:

$$-\frac{\kappa}{T^2} \operatorname{Tr}\left(\log(\mu(T))^2\right) = 1 \tag{7.7}$$

where the matrix μ is defined as:

$$\mu(T) := \begin{pmatrix} \frac{\sin(TB)(B^y + iB^x)}{B} & -\cos(TB) \\ \cos(TB) & \frac{\sin(TB)(B^y - iB^x)}{B} \end{pmatrix} \tag{7.8}$$

for convenience.

The 7.7 for the optimal time can be numerically solved in order to obtain the least positive root T_{opt} which is equal to the desired optimal time. The trace appearing in 7.7 (due to the cyclic property of the matrix trace) depends on the eigenvalues of the quantity within the trace. Thus one should diagonalise before attempting to numerically solve 6.36. It greatly simplifies the process and can be easily achieved with any good algebra package. The two eigenvalues of the matrix μ are:

$$\lambda_\pm(T) = B^y \sin(TB) \pm \frac{1}{\sqrt{2}}\left(\sqrt{((B^x)^2 + 1)\cos(2TB) - (B^x)^2 + 1}\right) \tag{7.9}$$

and thus the required time T_{opt} is the least root to the equation:

$$T^2 - \kappa\left(\log(\lambda_+(T))^2 + \log(\lambda_-(T))^2\right) = 0 \tag{7.10}$$

Here the fact that the trace of a normal matrix is the sum of its eigenvalue has been used. The roots of this equation, and thus the desired optimal times, can be found numerically by using a standard root finding algorithm. The optimal times obtained from this equation warrant direct comparison to 7.10.

For a concrete example we fix $\kappa = \frac{1}{2}$, i.e. $D = 1$. The optimal times for the complete range of values of B which meet the premise that $\operatorname{Tr}\left(\hat{H}_0^2\right) \leq 2$ i.e. $-\frac{1}{\sqrt{2}} \leq B \leq -\frac{1}{\sqrt{2}}$ are shown in figure 7.2. These times were obtained by using a Newton-Raphson root finding method with starting point $T = 1$. We found that the computation time for these optimal times 7.2 were negligible on a standard desktop computer using only a single core.

In the time independent case, the times can be obtained by substituting into 7.6: Then the time in the time independent case is:

$$T_{\text{opt}} = \frac{\pi}{2}\frac{B}{(1 - 2B^2)}\left(1 \pm \frac{\sqrt{1 - B^2}}{B}\right) \tag{7.11}$$

This function is also plotted in figure 7.2 for comparison to the optimal times in the time dependent case.

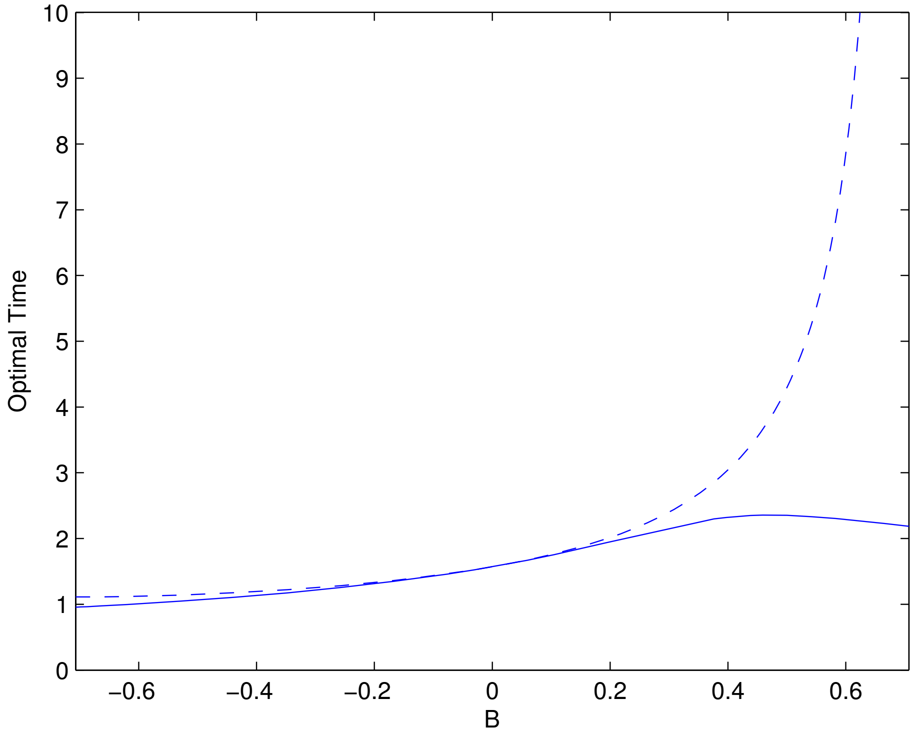

Figure 7.1: Times for time dependent (solid) and independent (dashed) controls.

7.2.1 Geodesics and Optimal Hamiltonian

By applying the first part of 6.33 one can obtain the time optimal trajectories for $\hat{U}_t$:

$$\hat{U}_t = \exp(-it\hat{H}_0)\exp\left(\frac{t}{T_{\text{opt}}}\log\left(\exp(iT_{\text{opt}}\hat{H}_0)\hat{O}\right)\right) \tag{7.12}$$

$$= \exp(-it\left(B^x\sigma_x + B^y\sigma_y\right))\exp\left(\frac{t}{T_{\text{opt}}}\log\left(\exp(iT_{\text{opt}}\left(B^x\sigma_x + B^y\sigma_y\right))\hat{O}\right)\right) \tag{7.13}$$

$$= \exp(-it\left(B^x\sigma_x + B^y\sigma_y\right))\exp\left(\frac{t}{T_{\text{opt}}}\log\left(\mu(T_{opt})\right)\right)$$

By applying the second part of 6.33 one can further obtain the optimal Hamiltonian driving $\hat{U}_t$ along such a geodesic: The optimal control Hamiltonian in closed form is:

$$\hat{H}_c(t) = \frac{i}{T_{\text{opt}}}\exp(-it\left(B^x\sigma_x + B^y\sigma_y\right))\log(\mu(T_{\text{opt}}))\exp(it\left(B^x\sigma_x + B^y\sigma_y\right))$$

In both of these formulae one has:

$$\exp(-it\hat{H}_0) = \tag{7.14}$$

$$\exp(-it\left(B^x\sigma_x + B^y\sigma_y\right)) = \begin{pmatrix} \cos(tB) & \frac{-\sin(tB)(B^y + iB^x)}{B} \\ \frac{\sin(tB)(B^y - iB^x)}{B} & \cos(tB) \end{pmatrix}$$

Unfortunately, further expanding the the other matrix exponentials and logarithms appearing seems not to lead to tractable expression.

7.2.2 Example of Extracting the Optimal Control Fields

The optimal control fields can also be determined for arbitrary positive value of $\kappa \in \mathbb{R}$. Consider that the control Hamiltonian can be expressed as $i\hat{H}_c(t) = D_t^k i\sigma_k$. One can now extract the individual control fields D^k from the following, as $i\sigma_k$ are an orthogonal basis (according to $h(i\hat{A}, i\hat{B}) = \kappa \operatorname{Tr}(\hat{A}\hat{B})$) for $\mathfrak{su}(n)$:

$$D_t^k = h\left(i\hat{H}_c(t), \frac{i\sigma_k}{2}\right) = \frac{\kappa}{2}\operatorname{Tr}(\hat{H}_c(t)\sigma_k) \tag{7.15}$$

Here the $\frac{1}{2}$ is included to ensure that $\frac{1}{2}i\sigma_k$ is an h unit vector.

As a concrete example of obtaining control fields, we set $B^x = B^y = \frac{1}{4}$. We solved the specific instance of 6.36 using the same root finding algorithm. The smallest, real, positive root is $T_{\text{opt}} \approx 3.2043\ldots$. (The actual physical time in seconds can be obtained by reintroducing the physical constants that have been lost after non-dimensionalisation throughout. Specifically, $\hbar$ has been set to 1 throughout.) In this instance the optimal control Hamiltonian is given by:

$$\hat{H}_c(t) = \tag{7.16}$$
$$\frac{i}{T_{\text{opt}}}\exp\left(\frac{-it}{4}(\sigma_x+\sigma_y)\right)\log\begin{pmatrix} \frac{1+i}{\sqrt{2}}\sin\left(\frac{T_{\text{opt}}}{\sqrt{8}}\right) & -\cos\left(\frac{T_{\text{opt}}}{\sqrt{8}}\right) \\ \cos\left(\frac{T_{\text{opt}}}{\sqrt{8}}\right) & \frac{1-i}{\sqrt{2}}\sin\left(\frac{T_{\text{opt}}}{\sqrt{8}}\right)\end{pmatrix}\exp\left(\frac{it}{4}(\sigma_x+\sigma_y)\right)$$

by 6.33. Evaluating the logarithm exactly in closed form is possible, but the result is cumbersome and does not provide any physical insight, so it is omitted.

The optimal control fields are obtained from this result. The kth field is obtained by evaluating 7.15 $\frac{1}{2}\operatorname{Tr}(\hat{H}_c(t)\sigma_k)$ numerically, as shown in 7.2. As a check, it has been numerically confirmed that the control fields have the property that the sum of their squares is $1/2$ ($\forall t$), which the constraint on $\operatorname{Tr}(\hat{H}_c(t)^2)$ mandates.

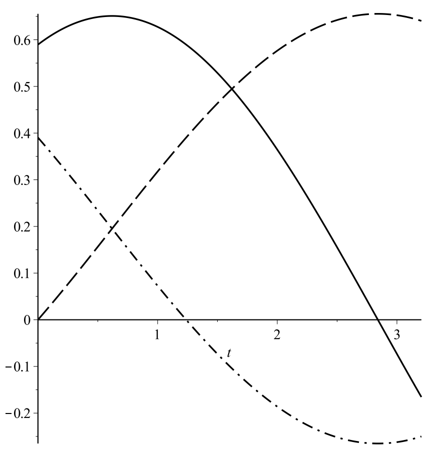

Figure 7.2: Optimal control fields D^x (dashed), D^y (dot-dashed), D^z (solid) for Pauli y gate, as a function of time, in the $B = 1/4$ case.

It is pleasing to observe that the fields driving the system along a geodesics are not too rapidly oscillating or degenerate (informally) in any other way. This indicates that it is plausible that such fields could be implemented in a practical laboratory context which validates the potential technological applications of the present work.

7.3 Two Coupled Spins: Constant vs Non-Constant Controls

For another, larger example, one can consider two coupled spins with Heisenberg coupling. The drift Hamiltonian for a two spin chain with (arbitrary spin coupling constants J_k) is [106]:

$$\hat{H}_0 = J^x \sigma_x \otimes \sigma_x + J^y \sigma_y \otimes \sigma_y + J^z \sigma_z \otimes \sigma_z \tag{7.17}$$

Again, we take h to be a multiple of the Killing form so that we can apply 6.3.1 to obtain the geodesics in closed form. By 6.33, the geodesic which connects $\hat{I}$ to $\hat{O}$ of the relevant Randers metric on $SU(4)$ are given by 6.37. One can exploit the block diagonal form of $\hat{H}_0$ in order to simplify the equation that needs to be solved.

By 6.33, the optimal Hamiltonian is:

$$\begin{aligned} \hat{H}_c(t) =& \frac{i}{T_{\text{opt}}} \exp\left(-it\left(J^x \sigma_x \otimes \sigma_x + J^y \sigma_y \otimes \sigma_y + J^z \sigma_z \otimes \sigma_z\right)\right) \\ & \times \log\left(\exp\left(iT_{\text{opt}}\left(J^x \sigma_x \otimes \sigma_x + J^y \sigma_y \otimes \sigma_y + J^z \sigma_z \otimes \sigma_z\right)\right)\hat{O}\right) \\ & \times \exp(it\left(\left(J^x \sigma_x \otimes \sigma_x + J^y \sigma_y \otimes \sigma_y + J^z \sigma_z \otimes \sigma_z\right)\right)) \end{aligned} \tag{7.18}$$

The actual control fields can be extracted, similarly to before, via:

$$f_{mn}(t) = \frac{\kappa}{4} \operatorname{Tr}\left(\hat{H}_c(t)\sigma_m \otimes \sigma_n\right) \tag{7.19}$$

This is the expansion of $\hat{H}_c(t)$ in a basis for $\mathfrak{su}(4)$. This basis is:

$$\{i\sigma_n \otimes \sigma_m \| n, m = 0, x, y, z \text{ but not both } n = 0 \text{ and } m = 0\} \tag{7.20}$$

Here σ^0 is taken to be the 2×2 identity matrix whereas the other σs are all the standard Pauli matrices. One readily checks that this basis is orthogonal w.r.t. the Killing form, which is the key property applied when extracting the control fields in 7.19. The origin of the 4 in this formula is the trace of the 4×4 identity matrix. Explicitly, we are representing $\hat{H}_c(t)$ as:

$$\hat{H}_c(t) = \sum f_{mn}(t)\sigma_m \otimes \sigma_n \tag{7.21}$$

where ths sum is over the basis vectors appearing in 7.20.

7.3.1 Explicit Calculation: The XXX-Chain

In order to again illustrate the way in which out method allows us to determine which systems are best suited to quickly implementing a QIP task, we study the case of the Isotropic Heisenberg spin chain, the $J^x = J^y = J^z = J$ case of 7.17. This leaves only one parameter J to consider, yielding a simple pedagogic example for the method. We set $\kappa = 1$ for the sake of example and thus consider the time for implementing the (special unitary) swap gate:

$$\hat{O} = e^{i\pi/4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{7.22}$$

Using 6.33, the optimal Hamiltonian is:

$$\begin{aligned} \hat{H}_c(t) = \frac{i}{4T_{\text{opt}}} & \begin{pmatrix} 2e^{-itJ} & 0 & 0 & 0 \\ 0 & e^{-itJ} + e^{3itJ} & e^{-itJ} - e^{3itJ} & 0 \\ 0 & e^{-itJ} - e^{3itJ} & e^{-itJ} + e^{3itJ} & 0 \\ 0 & 0 & 0 & 2e^{-itJ} \end{pmatrix} \\ & \times \log \left(\frac{1}{2} e^{\frac{i\pi}{4}} \begin{pmatrix} 2e^{iT_{\text{opt}}J} & 0 & 0 & 0 \\ 0 & e^{iJT_{\text{opt}}} - e^{-3iJT_{\text{opt}}} & e^{iJT_{\text{opt}}} + e^{-3iJT_{\text{opt}}} & 0 \\ 0 & e^{iJT_{\text{opt}}} + e^{-3iJT_{\text{opt}}} & e^{iJT_{\text{opt}}} - e^{-3iJT_{\text{opt}}} & 0 \\ 0 & 0 & 0 & 2e^{iJT_{\text{opt}}} \end{pmatrix} \right) \\ & \times \begin{pmatrix} 2e^{itJ} & 0 & 0 & 0 \\ 0 & e^{itJ} + e^{-3iJt} & e^{itJ} - e^{-3itJ} & 0 \\ 0 & e^{itJ} - e^{-3iJt} & e^{itJ} + e^{-3itJ} & 0 \\ 0 & 0 & 0 & 2e^{itJ} \end{pmatrix} \end{aligned} \tag{7.23}$$

We can determine the optimal time T_{opt} as before, by numerically solving 6.36. Substituting the specifics of the current problem into this equation yields:

$$-\frac{1}{T^2} \operatorname{Tr} \left(\log \left(\frac{1}{2} e^{\frac{i\pi}{4}} \begin{pmatrix} 2e^{iTJ} & 0 & 0 & 0 \\ 0 & e^{iJT} - e^{-3iJT} & e^{iJT} + e^{-3iJT} & 0 \\ 0 & e^{iJT} + e^{-3iJT} & e^{iJT} - e^{-3iJT} & 0 \\ 0 & 0 & 0 & 2e^{iJT} \end{pmatrix} \right) \right)^2 = 1$$

Using the same numerical procedure as in the previous example, we obtain the optimal execution times. These are shown in 7.3.

The dashed line in 7.3 plots the function, adapted from [2] to this specific scenario, giving the optimal time obtainable using only time independent control fields:

$$T_{opt} = \frac{\pi}{2} \left(\frac{\sqrt{3}}{2\sqrt{3}J \pm 1} \right) \tag{7.24}$$

where the $\pm$ is again chosen to make the time positive.

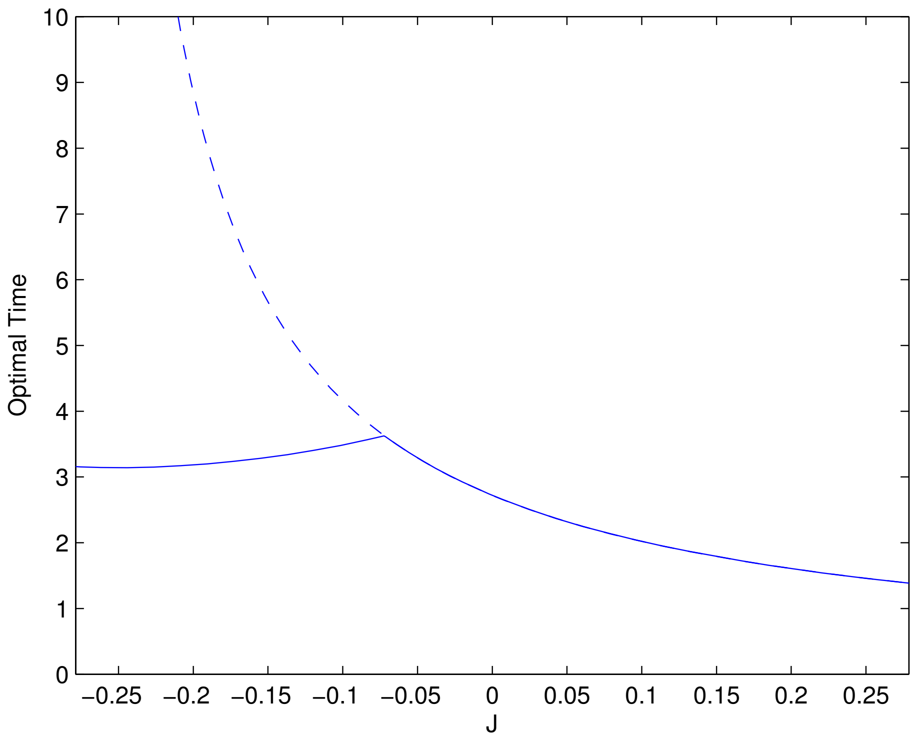

Figure 7.3: Times for time dependent (solid) and time independent (dashed) controls in an XXX Spin Chain.

7.4 Analysis of Results

7.4.1 Homogeneous Geodesics

A *homogeneous geodesic* (through the identity) on a Lie group with a Finsler metric F is a one-parameter subgroup which is also a geodesic [107, 108]. Here, it is a curve of the form $\hat{U}_t = \exp(-it\hat{A})$ for some $i\hat{A} \in \mathfrak{su}(n)$, which is also a geodesic of F. The $-i\hat{A} \in \mathfrak{su}(n)$ is a called a *geodesic vector*. These are exactly the curves that can be trajectories of a controlled quantum systems (of the type discussed throughout) for which only constant controls are permitted, as discussed in [2].

Theorem 3.1 in [108] presents a condition, which needs to be mildly adapted to apply to the scenario of Zermelo navigation on $SU(n)$. Adapted to the present example of $SU(n)$ (rather than a general homogeneous space as in the original work), the condition for $\hat{X}$ to be a geodesic vector is:

$$g_{\hat{X}}(\hat{X}, [\hat{X}, \hat{Z}]) = 0, \ \forall \hat{Z} \in \mathfrak{su}(n) \tag{7.25}$$

where g is the fundamental tensor of F restricted to $\mathfrak{su}(n)$. Here the fundamental tensor of a Finsler metric F on $SU(n)$ is the Hessian of the same metric point-wise (within a single tangent space):

$$g_{\hat{A}}(\hat{B}, \hat{C}) = \frac{1}{2}\frac{\partial^2}{\partial s \partial t}\left\{F^2\left(\hat{A} + s\hat{B} + t\hat{C}\right)\right\} \tag{7.26}$$

In the specific case of F a Randers metric, the fundamental tensor is given by formula (1.5) in [23].

Many of the results about homogeneous geodesics of Finsler metrics on Lie groups are applicable to *left* invariant metrics. However, these results can be easily adapted to the right invariant case, which arises naturally in quantum mechanics. The construction of an "opposite group" allows one to adapt results without difficulty; typically only sign changes are incurred.

7.4.2 A Conjecture

Theorem 3.7 in [108] can be applied to establish that any right invariant Randers metric on $SU(n)$ $(n \geq 2)$ will have infinitely many homogeneous geodesics. The theorem establishes that a right invariant Finsler metric on a compact Lie group will have infinitely many geodesic vectors. This can be easily seen by direct application of the theorem and by the observation that the rank of $SU(n)$ is $n-1$, and thus the theorem applies to any qubit system of more than one qubit. The theorem requires that the Lie group rank is ≥ 2, thus in all cases except $SU(2)$ (the single qubit case for which the rank is 1), there exist infinitely many homogeneous geodesics. This establishes the importance of determining all homogeneous geodesics of right invariant Randers metrics on $SU(n)$. Furthermore, in the case of a Randers metrics (as is our case) theorem (4.2) in [108] establishes a practically simplifying condition on homogeneous geodesics.

The author conjectures that this theorem is the underlying explanation for the two curves in figure (7.3) being equal for a continuous region of J values is that, in this region, the time independent trajectories are homogeneous geodesics of the Randers metric F. Confirming this conjecture will require further work.

Chapter 8

Additional Constraints and Sub-Finsler Metrics On $SU(n)$

This chapter covers methods for handling additional constraints on $\hat{H}_c$ in the first problem of quantum control. This work was first published in [3].

8.1 The Physical Need For Additional Constraints

The constraints studied thus far are, alone, insufficient for many physical applications. The assumption thus far as been only that the control Hamiltonian is constrained to be such that $h(i\hat{H}_c(t), i\hat{H}_c(t)) = 1$. In comparison to [4], only the roles of the L_T and L_S parts of the Lagrangian have been treated here. The L_T part is analogous to our application of the result of [29] on Zermelo navigation in the case of a right invariant Riemannian metric. The L_S part has no analogue as our work expresses the problem in a geometrically intrinsic way. We consider this to be an advantage of our method as it allows a more mathematical view of the problem; intrinsic geometry has been proven many times to be superior for mathematical analysis of geometric problems over methods based on many constraints or specific coordinate systems. This allows us to formulate the problem as a *first order* system of ODE from the outset by using the EP equation, rather than needing to compute any first variations or use the EL equations. So we can obtain a first order equation for the optimal Hamiltonian, *without* the need to actually determine any geodesics a priori.

There is also a disadvantage of the methods presented in this thesis compared with [4]: they can handle fewer types of constraints. Our method thus far can handle only the cases where the 'size' type constraint [4] is representable by an inner product.

8.1.1 Motivating Example: Heisenberg Spin Chains

To motivate the need for further constraints, we again consider the drift Hamiltonian for a two spin "chain" (with anisotropic couplings J) [106]:

$$\hat{H}_0 = J_x \sigma^x \otimes \sigma^x + J_y \sigma^y \otimes \sigma^y + J_z \sigma^z \otimes \sigma^z \tag{8.1}$$

Simply constraining the control Hamiltonian to be such that $h(i\hat{H}_c(t), i\hat{H}_c(t)) = 1$ for some inner product h is insufficient for practical applications where the producible set of control Hamiltonians does not include every direction within $\mathfrak{su}(n)$. For example, a common model of a controlled spin chain is one in which the control Hamiltonian takes the form:

$$\hat{H}_c(t) = f_1(t)\sigma_z \otimes \hat{I} + f_2(t)\hat{I} \otimes \sigma_z \tag{8.2}$$

That is, there is one local control field in the z direction only for each site in the chain. In such a situation no terms like $\sigma^x \otimes \sigma^x$ (or multiples there of) could appear in the control Hamiltonian, as these represent the couplings between sites in the chain, and are not the effect of *any possible* external field. In this case (choosing h to be κ times the Killing form), the constraint $h(i\hat{H}_c(t), i\hat{H}_c(t)) = 1$ evaluates to $\kappa \operatorname{Tr}(\hat{H}_c(t)^2) = 1$, which only constrains the sum of the squares of the control fields. An extra constraint must be added to exclude those terms from the control Hamiltonian that cannot be physically implemented.

8.1.2 Implementing Constraints Using Lagrange Multipliers

Additional constraints can be implemented by including Lagrange multipliers to create a new functional:

$$\Lambda\left(\hat{U}_t, \frac{d\hat{U}_t}{dt}, \omega_k(t)\right) = \frac{1}{2}F_{\hat{U}_t}\left(\frac{d\hat{U}_t}{dt}\right)^2 + \sum_k \omega_k(t)\left(f_{k,\hat{U}_t}\left(\frac{d\hat{U}_t}{dt}\right) - c_k\right) \tag{8.3}$$

where F is the Randers metric solving the relevant navigation problem, f_k represent the additional constraints, and ω_k are the Lagrange multipliers. The values c_k represent the value of f to which the trajectory is constrained.

We consider only the case where f is right invariant; this results in Λ also being right invariant. This corresponds to situations where the additional constraints depend only on the Hamiltonian, rather than the on location of $\hat{U}_t$ on $SU(n)$.

In this situation Λ can be expressed as:

$$\Lambda\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger, \omega_k(t)\right) = \frac{1}{2}F\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger\right)^2 + \sum_k \omega_k(t)\left(f_k\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger\right) - c_k\right) \tag{8.4}$$

$$= \frac{1}{2}F\left(-i\hat{H}(t)\right)^2 + \sum_k \omega_k(t)\left(f_k\left(-i\hat{H}(t)\right) - c_k\right) \tag{8.5}$$

where $\hat{H}_t$ is the Hamiltonian such that $\hat{U}_t$ solves the corresponding Schrödinger equation.

8.1.3 'Forbidden Directions'

One specific set of f_k and c_k with practical relevance is:

$$f_k\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger\right) = \operatorname{Tr}\left(\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger + i\hat{H}_0\right)i\hat{F}_k\right). \tag{8.6}$$

with $c_k = 0$.

This corresponds to the $\hat{F}_k$ spanning a set of "forbidden" terms for the control Hamiltonian. One can check this interpretation of the constraint by noticing that if $\hat{U}_t$ solves the Schrödinger equation with a Hamiltonian of the form of 4.6, then variation of 8.4 by ω_k yields:

$$\mathrm{Tr}\left(\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger + i\hat{H}_0\right) i\hat{F}_k\right) \tag{8.7}$$

which implies:

$$\mathrm{Tr}\left(\hat{H}_c(t)\hat{F}_k\right) = 0$$

and thus the desired "forbidden" directions are trace-orthogonal to the control Hamiltonian, and the control Hamiltonian has no component in any forbidden direction. These are essentially identical to the "linear homogeneous" constraints in [4]. There is a subtle difference however: here the forbidden direction applies only to the control Hamiltonian and not the overall Hamiltonian. This constraint is equivalent to an *affine* constraint on the overall Hamiltonian. Adding too many additional constraints may render the system in question uncontrollable. Existence/uniqueness of optimal trajectories is an issue not addressed in [4].

In order to find the equation satisfied by the optimal Hamiltonian that takes into account some additional constraints, we must modify 6.19. In the remainder of this paper we consider only the "forbidden direction" type of additional constraint. The equations satisfied by the optimal Hamiltonian (if any exist) can be found (in a basis for $\mathfrak{su}(n)$) by variation of each dependent variable on which Λ depends. Variation by $\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger$ yields the Euler-Poincaré equation for Λ:

$$\frac{d}{dt}\frac{\partial\Lambda}{\partial\xi^d} = -C^a_{bd}\frac{\partial\Lambda}{\partial\xi^a}\xi^b \tag{8.8}$$

Variation by ω_k yields:

$$\mathrm{Tr}\left(\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger + i\hat{H}_0\right) i\hat{F}_k\right) = 0 \tag{8.9}$$

Equations 8.8 and 8.9 need to be solved simultaneously in order to obtain the optimal Hamiltonian.

Obtaining closed form solutions for these equations and further numerical solution techniques in specific cases of physical interest requires further work. We would need to perform a complete analysis of common two qubit gates implemented in spin chain systems and other laser driven models. We would then hope to obtain an exact formula for the initial conditions required (for the system 8.8 and 8.9) in order for numerical solution of the system of ODEs to yield a geodesic connecting $\hat{I}$ to an arbitrary desired gate $\hat{O}$. A method for achieving a very similar goal has be found in a very different context [109]; only the Riemannian case is addressed, but it seems that the technique is easily adaptable.

The Role of Sub-Finsler Geometry

Mathematically, the optimal trajectories can be understood as sub-Finsler geodesics, or more specifically, what could be appropriately called sub-Randers geodesics. Many of the relevant definitions in this section can for found here [110] where they are discussed in great detail. One can find information about sub-Riemannian geometry in optimal control in [111]. For an application of sub-Finsler geometry in optimal control see [112]. For a specific application in quantum control see [113].

Definition 8.1.1. A *Distribution* on a manifold M is a smooth assignment of subspaces $V_p \subset T_pM$ of fixed dimension ($\dim(V_p) = k \leq \dim(M)$). Such an assignment can be specified by choosing k linearly independent (point wise) vector fields $V_p^{(k)}$.

Definition 8.1.2. A distribution $\left\{V_p^{(k)}\right\}$ on M is called *Bracket Generating* iff every smooth vector field $X \in \Gamma[TM]$ can be expressed in as a linear combination of iterated commutators of the vector fields $\left\{V_p^{(k)}\right\}$.

The now well known "Hormander's condition" [114] for the controllability of an affine control system on a Lie group of the type studied in this work, that is systems of the form of 4.2, indicates which sets of forbidden direction results in the system no longer being controllable. In the quantum case the vector fields spanning the distribution are all right invariant and the bracket generating condition corresponds to a condition that holds at a single point on the group $SU(n)$. These vector fields in the case of a controlled quantum system of the form 4.6 are:

$$\mathcal{S} = \left\{-i\hat{H}_0, -i\hat{H}_1, \ldots, -i\hat{H}_N\right\} \tag{8.10}$$

where $\hat{H}_0$ is the drift Hamiltonian and the others are a basis for the possible control Hamiltonians. The condition for controlability in the quantum case is that the Lie algebra Lie(S) bracket generated by the set S satisfies Lie$(S) = \mathfrak{su}(n)$. If this condition fails then this would mean that there were unitary gates that could not be implemented using a system constrained in such a way. It is however worth noting that this condition makes no reference to the control fields having restricted size. This adds additional restrictions to those studied in Hormander's condition and may still result in uncontrollability despite Hormander's condition being met. In short, in the presence of a constraint on the size of the control fields, Hormander's condition is merely necessary for controlability but is not sufficient.

In order to fully describe the geometry of quantum systems where there is an uncontrollable drift Hamiltonian a definition is needed:

Definition 8.1.3. An *Affine Distribution* on a manifold M is a smooth assignment of affine subspaces $V_p \subset T_pM$ of fixed dimension ($\dim(V_p) = k \leq \dim(M)$). Such an assignment can be specified by choosing k linearly independent (point wise) vector fields $V_p^{(k)}$ and another vector field U_p. The points of the distribution as a subset $\mathcal{D}_p$ of T_pM are given by:

$$\mathcal{D}_p = U_p + \text{Span}\left\{V_p^{(k)}\right\} \tag{8.11}$$

The method for handling the constraint $h(i\hat{H}_c(t), i\hat{H}_c(t)) = 1$, alongside additional constraints, shows that the optimal trajectories for $\hat{U}_t$ are geodesics of Randers metric parallel to an *affine* distribution, say $\mathcal{D}$, on $SU(n)$. $\mathcal{D}$ is the distribution consisting of vectors in $TSU(n)$ of the form:

$$\mathcal{D}_{\hat{U}} = -i\hat{H}_0\hat{U} + \text{Span}\{i\hat{H}_k\hat{U}\big|k = 0, \dots, N\} \tag{8.12}$$

Here $\{i\hat{H}_k\big|k = 0, \dots, N\} \subset \mathfrak{su}(n)$ span the subset of $\mathfrak{su}(n)$ that is h-orthogonal to the span of the subset of $\mathfrak{su}(n)$ spanning the "forbidden directions". This distribution is right invariant in the sense that: $\mathcal{D}_{\hat{U}} = \mathcal{D}\hat{U}$. That is, the optimal trajectories are the length minimising curves that connect given endpoints ($\hat{I}$ to $\hat{O}$) according a Randers metric F (solving the navigation problem in our case), and which are parallel to the distribution $\mathcal{D}$.

Definition 8.1.4. A curve $\hat{V}_t$ is called *parallel* to a (potentially affine) distribution $\mathcal{D}$ iff $\frac{d\hat{V}_t}{dt} \in \mathcal{D}_{\hat{V}_t}$ $\forall t$.

One can see that the condition that a curve on $SU(n)$ satisfies the EP equation with additional constrains described in this chapter is exactly the condition to be a geodesic of a Randers metric parallel to an affine distribution. 8.8 imposes that a curve is an, at least local, extremal curve of the length functional. 8.9 can then be understood as imposing the curve is parallel according to $\mathcal{D}$ as defined in 8.12. The system formed by 8.8 and 8.9 are solved by such curves.

Such curves are definitively sub-Randers geodesics. Thus it has been shown that the problem of quantum time optimal control for systems meeting the premises described in this chapters can be mathematically stated as the problem of finding sub-Randers geodesics for a right invariant sub-Randers metric. The right invariance follows from the right invariance of the navigation metric and the distribution $\hat{D}_{\hat{U}}$.

Chapter 9

Example Calculations With Additional Constraints

This chapter covers example calculations for handling additional constraints on $\hat{H}_c$ in the first problem of quantum control. This work was first published in [3].

9.1 Single Spin With One Forbidden Direction

Here it is illustrated how a forbidden direction can be treated in the example of a single spin. For simplicity, we consider the case that there is no drift. This makes the navigation metric F Riemannian, which makes ℓ a quadratic function. This allows us to solve for the time derivative of ξ explicitly in the EP equations, and then integrate the equation in closed form by hand. The case with drift is conceptually identical, except it may not always be possible to solve for $\dot{\xi}$, so the resulting system is more difficult to solve analytically and would require a numerical solution.

A system with control Hamiltonian constrained such that:

$$\frac{1}{2}\operatorname{Tr}(\hat{H}_c(t)^2) = 1 \tag{9.1}$$

is considered. Writing $i\hat{H}_c(t) = \xi^k(t) i\sigma_k$ one sees that this condition is $\xi^{x2}+\xi^{y2}+\xi^{z2} = 1$. Suppose further that one restricts to $\xi^z = c$ for $\forall t$. This is different from the examples worked out in [5], as this is an affine constraint rather than a linear homogeneous constraint. In a situation with a drift term, it is simple to see that a linear constraint on the control Hamiltonian corresponds to an equivalent affine constraint on the overall Hamiltonian.

In the present case the overall Lagrangian is:

$$\Lambda(\hat{H}_t, \omega(t)) = \frac{1}{2}\left(\xi^{x2} + \xi^{y2} + \xi^{z2}\right) + \omega(t)\left(\xi^z - c\right) \tag{9.2}$$

The EP equations are:

$$\begin{pmatrix} \dot{\xi^x} \\ \dot{\xi^y} \\ \dot{\xi^z} + \dot{\omega} \end{pmatrix} = \begin{pmatrix} -\omega\xi^y \\ \omega\xi^x \\ 0 \end{pmatrix} \tag{9.3}$$

and $\xi^z(t) = c$, $\dot{\xi^z} = 0$. which implies:

$$\begin{pmatrix} \dot{\xi^x} \\ \dot{\xi^y} \end{pmatrix} = \begin{pmatrix} -\omega\xi^y \\ \omega\xi^x \end{pmatrix} \tag{9.4}$$

and that ω is a constant. The general solution (after imposing the unit speed condition) is:

$$\begin{aligned} \xi^x(t) &= A\cos(\omega t) - \sqrt{1-c^2-A^2}\sin(\omega t) \\ \xi^y(t) &= A\sin(\omega t) + \sqrt{1-c^2-A^2}\cos(\omega t) \\ \xi^z(t) &= c \end{aligned} \tag{9.5}$$

wherein A is an arbitrary constant parameter and ω is the Lagrange multiplier. The trajectories in $\mathfrak{su}(n)$ are circles in $\mathfrak{su}(2)$ centred at:

$$\begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix} \tag{9.6}$$

The unit speed condition imposed is ensure that the parameter in $\hat{H}_t$ is physical time as it appears in the Schrödinger equation.

Both A and ω parametrise possible endpoints of $\hat{U}_t$ after the trajectory on the group is reconstructed via the Schrödinger equation. The optimal Hamiltonians are:

$$\hat{H}_t = (A\cos(\omega t) - B\sin(\omega t))\sigma_x + (A\sin(\omega t) + B\cos(\omega t))\sigma_y + c\sigma_z \tag{9.7}$$

where $B = \pm\sqrt{1-c^2-A^2}$.

9.2 Single Spin With One Forbidden Direction and Drift

In order to illustrate the effect of adding a drift Hamiltonian on the navigation metric, set:

$$i\hat{H}_0 = \frac{1}{2} i\sigma_x \tag{9.8}$$

for some constant vector χ. As before: $i\hat{H}_c(t) = \xi^k(t) i\sigma_k$ and the same constraint is imposed.

$$\begin{aligned} \Lambda(\hat{H}_t, \omega(t)) &= \frac{1}{2} F\left(\xi^k i\sigma_k\right)^2 + \omega(t)\left(\frac{1}{2}\operatorname{Tr}\left(\xi^k\sigma_k\sigma_x\right) - c\right) \\ &= \frac{1}{2}\left(\sqrt{\frac{2}{3}\left(\xi^{y2}+\xi^{z2}\right) + \frac{10}{9}\left(\xi^{x2}\right)} - \left(\frac{2}{3}\xi^{x2}\right)\right)^2 + \omega(t)\left(\xi^x - x\right) \\ &= -\frac{2}{3}\xi^{z2}\sqrt{\frac{10}{9}\xi^{x2} + \frac{2}{3}\left(\xi^{y2}+\xi^{z2}\right)} \\ &+ \frac{5}{9}\xi^{x2} + \frac{1}{3}\xi^{y2} + \frac{2}{9}\xi^{z4} + \frac{1}{3}\xi^{z2} + \omega(t)\left(\xi^x - x\right) \end{aligned} \tag{9.9}$$

9.3 Obstacles to Solving the General Case

The metric 6.5 has been obtained in two physically practical cases. The EP equation has been obtained for the drift free case.

It is possible to obtain the EP equation for the Lagrangian 9.9. However, as far as the author knows, it is not possible to solve these equations for $\dot{\xi}^k$. This is a consequence of the persistence of the square root term after expanding out the power of two in 9.9. The author feels that, at this point, one must resort to numerical methods.

This appears to be the generic case as the persistence of the square root term in the Lagrangian is a feature of the EP equation associated to any Randers norm and is not unique to the case 9.9. By considering the most general Randers norm on $\mathfrak{su}(n)$:

$$F(\hat{A}) = \sqrt{\alpha\left(\hat{A},\hat{A}\right)} + \beta\left(\hat{A}\right) \tag{9.10}$$

one finds:

$$\ell(\hat{A}) = \frac{1}{2}\left(\sqrt{\alpha\left(\hat{A},\hat{A}\right)} + \beta\left(\hat{A}\right)\right)^2 \tag{9.11}$$

$$= \frac{1}{2}\left(\alpha\left(\hat{A},\hat{A}\right) + 2\sqrt{\alpha\left(\hat{A},\hat{A}\right)}\beta\left(\hat{A}\right) + \beta\left(\hat{A}\right)^2\right) \tag{9.12}$$

$$= \frac{1}{2}\left(\alpha_{ij}\xi^i\xi^j + 2\sqrt{\alpha_{ij}\xi^i\xi^j}\beta_k\xi^k + (\beta_k\xi^k)^2\right)$$

wherein $\hat{A} = \xi^k\hat{G}_k$ and $\{\hat{G}_k\}$ are a basis for $\mathfrak{su}(n)$. This leads to the EP equation 6.19 for this Lagrangian.

As of yet, the author does not know a method for solving this equation. As the EP equation associated to any non-Riemannian right invariant Randers metric on $\mathfrak{su}(n)$ is an implicit (in $\dot{\xi}^k$) differtial equation, it is probably very difficult if not impossible to solve the general case in closed form.

Chapter 10

Time Optimality Equations

This section presents the most general form of the equations for time optimality implimenting a quantum gate in a constrained system. These equations are shown to be the geodesic equations for a sub-Finsler metric on $SU(N)$.

10.1 a Broad Generalisation

The method for handling forbidden directions can be generalised to include a much larger class of constraints replacing the constraint that $h(i\hat{H}_c(t).i\hat{H}_c(t)) = 1$. One can replace the role of h representing the constraint of the size of $i\hat{H}_c(t)$ with an arbitrary right invariant Finsler metric, which we denote by $\check{F}_{\hat{U}}$, i.e. we now, more generally than before, impose $\check{F}(i\hat{H}_c(t)) = 1$ holds at the identity on $SU(n)$. As right invariant Finsler metrics on $SU(n)$ are in one-to-one correspondence with Minkowski norms on $\mathfrak{su}(n)$, this new class of constraints is much larger that the class of right invariant Riemannian metrics employed before.

In this chapter Shen's [29] Lemma 3.1 adapted to the case of $SU(n)$ with a right invariant $\check{F}$. In addition to the exact solution for the Riemannian case given by Shen's theorem, this gives an equation for a Finlser metric F the geodesics of which are time optimal the the presence of the constraint $\check{F}(i\hat{H}_c(t)) = 1$.

$$\check{F}\left(\frac{\hat{A}}{F(\hat{A})} + i\hat{H}_0\right) = 1 \qquad \forall \hat{A} \in \mathfrak{su}(n)/\{0\} \tag{10.1}$$

Note that the roles of F and $\check{F}$ are reversed here compared with the original presentation. One can easily check that the solution for F will be a Randers metric exactly when $\check{F}$ is Riemannian; this is exactly the case solved by Shen's theorem.

The premise that the 'wind'/drift Hamiltonian can be overcome by the control is now $\check{F}(i\hat{H}_0) \leq 1$. This guarantees that the desired time optimal trajectories are the geodesics of the Finsler metric F solving (10.1). The solution F is right invariant if both $\check{F}$ and the drift vector field are, as is the case for quantum control problems. This follows directly from substituting right invariant $\check{F}$ and drift vector field into the equation defining F and then right translating to the identity.

Now we give the set of equation that define the optimal trajectories in such a scenario. Time optimality yields:

$$\frac{d}{dt}\frac{\partial \Lambda}{\partial \xi^d} = -C^a_{bd}\frac{\partial \Lambda}{\partial \xi^a}\xi^b \tag{10.2}$$

Variation by ω_k to impose the forbidden direction constraints, as before, yields:

$$\mathrm{Tr}\left(\hat{H}_c(t)\hat{F}_k\right) = 0 \tag{10.3}$$

Here Λ is as before, except the F is no longer necessarily a Randers metric, but is now the solution to (10.1). This solution is guaranteed to also be a Finsler metric [29].

10.2 The General Time Optimality Equation

Together this all yields the system for the time optimal Hamiltonian $\hat{H}_t = \hat{H}_0 + \hat{H}_c(t) = \xi^k \hat{G}_k$:

$$\begin{cases} \check{F}\left(\dfrac{\hat{A}}{F(\hat{A})} + i\hat{H}_0\right) = 1 \qquad \forall \hat{A} \in \mathfrak{su}(n)/\{0\} \\ \\ \dfrac{d}{dt}\dfrac{\partial \Lambda}{\partial \xi^d} = -C^a_{bd}\dfrac{\partial \Lambda}{\partial \xi^a}\xi^b \\ \\ \mathrm{Tr}\left(\hat{H}_c(t)\hat{F}_k\right) = 0 \qquad \forall k, \forall t \geq 0 \\ \\ \hat{U}_T = \mathcal{T}\exp\left(\displaystyle\int_0^T -i\hat{H}_t dt\right) = \hat{O} \end{cases} \tag{10.4}$$

These we refer to as the time optimality equations for the gate $\hat{O}$, the drift Hamiltonian $\hat{H}_0$ and the constraint that $\check{F}(\hat{H}_c(t)) = 1\ \forall t \geq 0$. Here, $\hat{G}_k$ are a basis for $\mathfrak{su}(n)$. These equations determine the optimal Hamiltonian.

As in [5], the author has not yet found a way to impose the boundary condition $\hat{U}_T = \hat{O}$ (for some T) without solving the other time optimality equations explicitly. It is, however, known which variations at the algebra level correspond to variations of $\hat{U}_t$ that leave the end points of a curve on $SU(n)$ fixed [92]. In quantum mechanical terms these are exactly variations of $-i\hat{H}_t$ of the form: $\delta i\hat{H}_t = i\frac{d\hat{K}_t}{dt} + [i\hat{H}_t, i\hat{K}_t]$. Here $\hat{K}_t$ is any smooth curve in $\mathfrak{su}(n)$ which is 0 at both end points. A method for imposing similar boundary conditions is presented in [109] in a different context. We hope to analyse that method and adapt it to quantum control scenarios, so that end point conditions on $\hat{U}_t$ can be imposed at the algebra level and thus the EP equations can still be applied.

Chapter 11

Speed Limits For State Control Problems

In this chapter state control problems are studied, where a given initial state $|\psi_I\rangle$ is driven to a given to a $|\psi_F\rangle$ in the least possible time. This task is compared to the problem studied in previous chapters of time optimally controlling the time evolution operator. The mathematical structure of quantum theory allows one to exploit the isomorphism described in 3.2.3 to translate between these two control problems.

The work of Brody and Meier on the quantum speed limit, which in part follows up to [2, 3], is also discussed. This work [115, 116] on the quantum speed limit for the navigation problem for quantum states was published during the writing of this thesis and some results were obtained at similar times independently and this interesting work is fully acknowledged.

Thus far, the issue of time optimally controlling $\hat{U}_t$ has been addressed exclusively. However, the work in [5] and many other papers address the issue of time optimally control a quantum state.

In this chapter, controlling the state of systems with a drift Hamiltonian $\hat{H}_0$ and limited size control fields are considered. The relationship between this problem and the problem of time optimally controlling $\hat{U}_t$ is elucidated.

11.1 Why Work With $\mathbb{C}P^n$ Rather Than $\mathbb{C}^{n+1}$ Directly

The state of a quantum system is typically described as a vector in a complex vector space. However, as described in 3.2, this description contains some redundancy.

This redundancy yields a reason to appeal to geometric descriptions of the quantum state as a point in a projective space in order to apply geometric control theory. In order to apply the formalism of Zermelo navigation to obtain optimal times one needs a manifold, a Riemanian metric and a drift field. $\mathbb{C}^{n+1}$ is both a vector space and a (real) manifold as its manifold structure is isomorphic to $\mathbb{R}^{2n+2}$. As such, one could consider Zermelo navigation on this manifold. However, there are obstacles to this approach which render $\mathbb{C}P^n$ a far superior setting for a navigation problem.

We will not work with the complex manifold structure of $\mathbb{C}^{n+1}$ or $\mathbb{C}P^n$ in this thesis, in fact $\mathbb{C}P^n$ is even a Kähler manifold when the Fubini Study metric is used. For a review of complex manifolds see [73], for interesting work on Zermelo navigation on complex manifolds see [117].

Two vectors $|\psi\rangle,\ Z|\psi\rangle \in \mathbb{C}^{n+1}$ represent the same physical state, where $Z \in \mathbb{C}/\{0\}$. Any curve $|\psi_t\rangle \in \mathbb{C}^{n+1}$ descends to a curve on $\mathbb{C}P^n$ given by $\gamma(|\psi_t\rangle) = [|\psi_t\rangle]$. Any new curve $Z_t|\psi_t\rangle \in \mathbb{C}^{n+1}$, where $Z_t \in \mathbb{C}/\{0\}$, descends to the same curve. In order to unambiguously define a metric structure on $\mathbb{C}P^n$ in terms of a function on $T\mathbb{C}^{n+1}$ one requires that: given a curve $|\psi_t\rangle$, all curves $Z_t|\psi_t\rangle \in \mathbb{C}^{n+1}$ are assigned the same length as they all descend to the same curve on $\mathbb{C}P^n$. This property is manifest in the definition of the Fubini-Study metric as a function on $T\mathbb{C}P^{n+1}$ given in 5.2.2. . This required invariance property of a function on $T\mathbb{C}^{n+1}$ means that it can never itself be a metric as it is degenerate, a property not possessed by either Riemanian or Finsler metrics. To illustrate this, consider the a given fixed normalised state $|\bar{\psi}\rangle \in \mathbb{C}^{n+1}$ and a curve $|\psi_t\rangle = Z_t|\bar{\psi}\rangle$. This curve as zero length according to the Fubini-Study metric, considered as a function on $T\mathbb{C}^{n+1}$. The tangent vector to the curve is simply: $|\delta\psi\rangle = \dot{Z}|\bar{\psi}\rangle$. This gives the length to be:

$$\begin{aligned} L[|\psi_t\rangle] &= \int_0^T \sqrt{\frac{\langle\delta\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle} - \frac{\langle\delta\psi|\psi\rangle\langle\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle^2}}dt \qquad (11.1)\\ &= \int_0^T \sqrt{\frac{|\dot{Z}_t|^2\langle\bar{\psi}|\bar{\psi}\rangle}{|Z_t|^2\langle\bar{\psi}|\bar{\psi}\rangle} - \frac{|\dot{Z}_t|^2|Z_t|^2\langle\bar{\psi}|\bar{\psi}\rangle^2}{|Z_t|^4\langle\bar{\psi}|\bar{\psi}\rangle^2}}dt \\ &= \int_0^T \sqrt{\frac{|\dot{Z}_t|^2}{|Z_t|^2} - \frac{|\dot{Z}_t|^2}{|Z_t|^2}}dt = \int_0^T 0\ dt = 0 \end{aligned}$$

This is not possible for any Finsler metric structure on $\mathbb{C}^{n+1}$ as the non-degeneracy condition for such a metric means that every (non-zero) tangent vector has non-zero, positive length. As such, any Finsler metric on $\mathbb{C}^{n+1}$ will assign a non-zero length to this curve, despite the curve representing a physical state remaining constant. Supposing that, as for any Finsler metric, $F_{|\psi\rangle}(|\delta\psi\rangle) > 0\ \forall|\psi\rangle \in \mathbb{C}^{n+1},\ \forall|\delta\psi\rangle \in T_{|\psi\rangle}/\{0\}$ one finds:

$$L[|\psi_t\rangle] = \int_0^T F_{|\psi\rangle}(|\delta\psi\rangle)\,dt > 0 \qquad (11.2)$$

In the Zermelo navigation formalism, the length of any curve (according to the navigation metri 6.5 gives the optimal time for traversing that curve. This analysis indicates that no Finlser metric structure on $\mathbb{C}^{n+1}$ can produce the physically correct answer, zero, for such optimal times as it cannot assign length 0 to the curve $Z_t|\bar{\psi}\rangle$. Physically speaking, un-physical changes in the phase (and length) of the quantum state vector are contributing to the length of the curve.

11.2 Zermelo Navigation On $\mathbb{C}P^n$

11.2.1 Setting Up the Navigation Problem

In order to set up a Zermelo navigation problem and again apply Shen's theorem, one requires a manifold, a Riemanian metric representing the a constraint on the control and a 'drift' vector field which is small according to the same metric. As described in 3.2, the state space in geometric quantum mechanics is a manifold $\mathbb{C}P^n$ and it is this manifold on which a navigation problem will be set up.

The Schrödinger equation of the time evolution of a quantum state is:

$$\frac{d}{dt}|\psi_t\rangle = -i\hat{H}_t|\psi_t\rangle \tag{11.3}$$

in the case of a driven system with drift Hamiltonian $\hat{H}_0$ this equation becomes:

$$\frac{d}{dt}|\psi_t\rangle = -i\left(\hat{H}_0 + \hat{H}_c\right)|\psi_t\rangle \tag{11.4}$$

Following a similar procedure to the case of Zermelo navigation on $SU(n)$, one can separate the right hand side of 11.4 in to two vector fields: $-i\hat{H}_0|\psi\rangle$ and $-i\hat{H}_c|\psi\rangle$. The first of these $-i\hat{H}_0|\psi\rangle$ is the drift field as it is time independent. Now a manifold and drift field are identified, what remains is to identify a Riemanian metric on $\mathbb{C}P^n$ in order to complete the ingredients for a Zermelo navigation problem.

11.2.2 Pull-Back Metrics On $SU(n+1)$

The next object required to set up a Zermelo navigation problem for the control of a quantum state is a metric on $\mathbb{C}P^N$. As expressed in 3.20, $\mathbb{C}P^n \cong SU(n+1)/U(n)$. As such we can obtain the required metric from a metric on $SU(n+1)/U(n)$. Furthermore, we can obtain a metric on $SU(n+1)/U(n)$ from a metric on $SU(n+1)$ as long as the appropriate invariance properties are satisfied. This sections develops these concepts specialised to the case of quantum mechanics, for a more general mathematically oriented expostulation of these concepts see [118].

The map $\phi : SU(n+1) \to \mathbb{C}P^n$ which achieves the difeomorphism in 3.20 is:

$$\phi(\hat{U}) = \hat{U} \circ [|\psi_0\rangle] = [\hat{U}|\psi_0\rangle] \tag{11.5}$$

where $|\psi_0\rangle$ is, as in 3.20, any state and $U(n)$ is the stabiliser of the equivalence class of this state. One readily checks that this map is 'constant on $U(n)$ cosets' in the following sense:

$$\phi\left(\hat{U}U(n)\right) = \left(\hat{U}U(n)\right) \circ [|\psi_0\rangle] = \hat{U} \circ (U(n) \circ [|\psi_0\rangle]) = \hat{U} \circ [|\psi_0\rangle] = \phi\left(\hat{U}\right) \tag{11.6}$$

where the penultimate step has applied the fact that $U(n)$ is the stabiliser of $[|\psi_0\rangle]$.

As such it is possible to unambiguously define a map $\chi : SU(n+1)/U(n) \to \mathbb{C}P^n$ by $\chi\left(\hat{U}U(n)\right) = \phi(\hat{U})$. The differential of the map $d\phi\big|_{[|\psi\rangle]} : T_{\hat{U}}SU(n+1) \to$

$T_{[\hat{U}|\psi_0\rangle]}\mathbb{C}P^n$. The relationship between the three manifolds can be represented by the following commutative diagram. If we define the quotient map $\pi : \mathbb{C}^{n+1} \to \mathbb{C}P^n$ by $\pi(|\psi\rangle) = [|\psi\rangle]$ then the following diagram commutes:

$$\begin{array}{ccc} SU(n+1) & \xrightarrow{\pi} & SU(n+1)/U(n) \\ & \searrow^{\phi} & \downarrow \chi \\ & & \mathbb{C}P^n \end{array}$$

The natural choice of group action, $\star : SU(n+1) \times \mathbb{C}P^n \to \mathbb{C}P^n$, of of $SU(n+1)$ on $\mathbb{C}P^n$ is $\hat{V} \circ [|\psi\rangle] := [\hat{V}|\psi\rangle]$. The sense in which this is the 'natural' choice of action is as follows.

$$\begin{array}{ccc} SU(n+1) \times \mathbb{C}^{n+1} & \xrightarrow{\diamond} & \mathbb{C}^{n+1} \\ {\scriptstyle \mathrm{id}\times\gamma}\downarrow & & \downarrow{\scriptstyle \gamma} \\ SU(n+1) \times \mathbb{C}P^n & \xrightarrow{\star} & \mathbb{C}P^n \end{array}$$

where γ is defined as in 3.2. This action allows one to consider quantum time evolution as happening on $\mathbb{C}P^n$ rather than in $\mathbb{C}^{n+1}$.

A Finlser metric F on $SU(n+1)$ which is right invariant, i.e. $F_{\hat{U}}\left(\hat{A}\hat{U}\right) = F_{\hat{I}}\left(\hat{A}\right)$, is manifestly constant on $U(n)$ cosets. A metric which is constant on $U(n)$ cosets can be pushed forward unambiguous through the projection map $\pi : SU\left(n+1\right) \to SU\left(n+1\right)/U\left(n\right)$. Similarly, a metric R on $\mathbb{C}P^n$ pulls back to a metric on $SU(n+1)$ which is constant on $U(n)$ cosets. This metric is defined by:

$$F(\hat{A}\hat{U}) := R_{\phi(\hat{U})}(d\phi|_{\hat{U}}(\hat{A}\hat{U})) \tag{11.7}$$

This discussion demonstrates that Finsler metrics on $\mathbb{C}P^n$, the space of quantum states, can always be pulled back to Finsler metrics $SU(n)$ which are constant on $U(n)$ cosets. In light of this Zermelo navigation problems on $\mathbb{C}P^n$ can always be converted into Zermelo navigation problems on $SU(n+1)$ for which the navigation metric is constant on $U(n)$ cosets. As the navigation metric, as shown in section, is always right invariant in the quantum context this is simply a special case of the problem already studied in previous chapters. Thus the optimal Hamiltonian achieving a state transfer problem can also be obtained by solving an Euler Poincaré equation in $\mathfrak{su}(n+1)$ as before.

11.2.3 Tangent Bundle of $\mathbb{C}P^n$

This section is included for clarity. In order to set up a Zermelo navigation problem on $\mathbb{C}P^n$ a drift vector field is required. Technically, $-i\hat{H}_0|\psi\rangle$ is a tangent vector on $\mathbb{C}^{n+1}$. The relationship between tangent vectors on $\mathbb{C}^{n+1}$ and $\mathbb{C}P^n$ is clarified here.

It is possible to express a tangent vector on $\mathbb{C}P^n$ in terms of a set of tangent vectors on $\mathbb{C}^{n+1}$. This can be achieved be differentiating $Z_t|\psi_t\rangle$ (with $|\psi_t\rangle$ normalized) in order to deduce which points in $T\mathbb{C}^{n+1}$ correspond to a single point in $T\mathbb{C}P^{n+1}$.

$$\frac{d}{dt}Z_t|\psi_t\rangle = \dot{Z}_t|\psi_t\rangle + Z_t|\delta\psi\rangle \tag{11.8}$$

as such one sees that any points in $T\mathbb{C}^{n+1}$:

$$\begin{aligned}&(|\psi\rangle, |\delta\psi\rangle) \\ &(Z|\psi\rangle, Y|\psi\rangle + Z|\delta\psi\rangle)\end{aligned} \tag{11.9}$$

$\forall Z \in \mathbb{C}/\{0\}$ and $Y \in \mathbb{C}$ correspond to a single point in $T\mathbb{C}P^n$. This indicates how $T\mathbb{C}P^n$ can be expressed as a quotient of $T\mathbb{C}^{n+1}$ in order to obtain a vector field on $\mathbb{C}P^n$ to use as the drift field.

11.3 Application To Orthogonality Times

The physical meaning of the connection between Zermelo navigation on $SU(n+1)$ and $\mathbb{C}P^n$ is that the optimal time to enact the transition from $[|\phi_I\rangle] \in \mathbb{C}P^n$ to $[|\phi_F\rangle] \in \mathbb{C}P^n$ is equal to the minimal time to drive $\hat{U}_t$ from the identity to an operator $\hat{O}$ achieving $\hat{O} \circ [|\psi_I\rangle] = [\hat{O}|\psi_I\rangle] = [|\psi_F\rangle]$.

Many bounds have been recently obtained for the minimum time for a system to transition between to orthogonal state, see 5.2.2 for a discussion of these. These bound predominantly apply to situations where the Hamiltonian is time independent and there is no uncontrollable drift field. First a theorem will be needed:

Theorem 11.3.1. *Given any special unitary operator $\hat{T}$ on $\mathbb{C}^N$ such that the following holds:*

- $\exists |\psi_0\rangle, |\psi_1\rangle \in \mathbb{C}^N$ *such that* $\hat{T}|\psi_0\rangle = |\psi_1\rangle$ *and* $\hat{T}|\psi_1\rangle = |\psi_0\rangle$
- $\langle\psi_1|\psi_0\rangle = 0$
- $\forall |\psi\rangle \in \{|\psi_0\rangle, |\psi_1\rangle\}^\perp$, $\hat{T}|\psi\rangle = |\psi\rangle$ *(where $\perp$ indicates the orthogonal complement)*

Then $\hat{T}$ has the form

$$\hat{T} = \hat{V}\hat{O}\hat{V}^\dagger \tag{11.10}$$

where

$$\hat{O} = \left(\begin{pmatrix} 0 & ie^{-i\theta} \\ ie^{i\theta} & 0 \end{pmatrix} \oplus \hat{I}_{N-2}\right) \tag{11.11}$$

for some special unitary $\hat{V}$.

That is, given any $\hat{T}$ that sends a specific pair of orthogonal states to each other, and leaves all other orthogonal states unchanged, then $\hat{T}$ can be expressed in a simplified form, of some unitary conjugation of the specific $\hat{O}$ given in the theorem. The operator $\hat{T} = \hat{V}\hat{O}\hat{V}^\dagger$ maps $\hat{V}|0\rangle$ to $\hat{V}|1\rangle$.

Proof Sketch. Choose $\hat{V}$ to be a change of basis matrix (all unitary matrices are change of basis matrices between orthonormal bases of $\mathbb{C}^N$). More specifically, take $\hat{V}$ to be a change from an orthonormal basis that includes $|\psi_0\rangle$ and $|\psi_1\rangle$ to a basis that includes $|0\rangle$ and $|1\rangle$. Choose a $\hat{V}$ that also satisfies $\hat{V}|\psi_0\rangle = |0\rangle$ and $\hat{V}|\psi_1\rangle = |1\rangle$. The theorem follows from this choice. □

Such a gate $\hat{T}$ will be referred to as an *Orthogonalising* gate.

11.3.1 A Two-Level Scenario

Consider a two level system without drift for which, as in section, the control Hamiltonian is constrained to be such that $\kappa \operatorname{Tr}(\hat{H}_c^2) = 1$. I.e. the constraint is represented by the bi-invariant metric. The optimal time to implement an orthogonalising gate $\hat{T}$ in a two level system is obtained.

Without Drift

In the scenario without drift the situation is simple as the navigation metric is simply the constraint metric $h(i\hat{A}, i\hat{A}) = \kappa \operatorname{Tr}((i\hat{A})^\dagger i\hat{A}) = \kappa \operatorname{Tr}(\hat{A}^2)$. As this is the bi-invariant metric on $SU(2)$, the geodesics, and thus the time optimal trajectories, are the one-parameter subgroups. These correspond to the trajectories of $\hat{U}_t$ obtained from constant control fields.

What remains is to find the length the geodesic connecting $\hat{I}$ to each $\hat{T}$ and then subsequently find the $\hat{T}$ for which this curve is shortest. The geodesic connecting $\hat{I}$ to $\hat{T} = \hat{V}\hat{O}\hat{V}^\dagger$, parametrised from $t = 0$ to $t = 1$, is:

$$\hat{U}_t = \exp\left(t \log(\hat{T})\right) = \exp\left(t \log(\hat{V}\hat{O}\hat{V}^\dagger)\right) \tag{11.12}$$

The generator of this one-parameter subgroup is:

$$\log(\hat{V}\hat{O}\hat{V}^\dagger) = \hat{V}\log(\hat{O})\hat{V}^\dagger = \frac{\pi}{2}\hat{V}\begin{pmatrix} 0 & ie^{-i\theta} \\ ie^{i\theta} & 0 \end{pmatrix}\hat{V}^\dagger \tag{11.13}$$

and thus the geodesics is:

$$\begin{aligned} \hat{U}_t &= \exp\left(t\frac{\pi}{2}\hat{V}\begin{pmatrix} 0 & ie^{-i\theta} \\ ie^{i\theta} & 0 \end{pmatrix}\hat{V}^\dagger\right) \\ &= \hat{V}\exp\left(\frac{t\pi}{2}\begin{pmatrix} 0 & ie^{-i\theta} \\ ie^{i\theta} & 0 \end{pmatrix}\right)\hat{V}^\dagger \end{aligned} \tag{11.14}$$

The length of this geodesics according to the navigation metric, which in the drift free case is simply the constraint metric is:

$$L[\hat{U}_t] = \sqrt{\kappa}\int_0^1 \mathrm{Tr}\left(\left\{\frac{\pi}{2}\hat{V}\begin{pmatrix}0 & ie^{-i\theta}\\ ie^{i\theta} & 0\end{pmatrix}\hat{V}^\dagger\right\}^\dagger\left\{\frac{\pi}{2}\hat{V}\begin{pmatrix}0 & ie^{-i\theta}\\ ie^{i\theta} & 0\end{pmatrix}\hat{V}^\dagger\right\}\right)^{1/2} dt \tag{11.15}$$

$$= \frac{\pi\sqrt{\kappa}}{2}\mathrm{Tr}\left(\begin{pmatrix}0 & ie^{-i\theta}\\ ie^{i\theta} & 0\end{pmatrix}^\dagger\begin{pmatrix}0 & ie^{-i\theta}\\ ie^{i\theta} & 0\end{pmatrix}\right) = \pi\sqrt{\kappa}$$

No optimisation is required here as $\hat{V}$ has dropped out of the calculation due to the bi-invariance of the metric.

With Drift

Consider a system with $\kappa = 1$ and drift Hamiltonian $\hat{H}_0 = B\sigma_x$ with $B \leq \frac{1}{2}$. The results of 6.3.2 show that the geodesics of the navigation metric associated to this scenario connecting $\hat{I}$ to $\hat{V}\hat{O}\hat{V}^\dagger$:

$$\hat{U}_t = \exp(-itB\sigma_x)\exp\left(\frac{t}{T}\log\left(\exp(iTB\sigma_x)\hat{V}\hat{O}\hat{V}^\dagger\right)\right) \tag{11.16}$$

and that the optimal time for implementing the gate is given by solving:

$$-\frac{\kappa}{T^2}\mathrm{Tr}\left(\log\left(\exp(iTB\sigma_x)\hat{V}\hat{O}\hat{V}^\dagger\right)^2\right) = 1 \tag{11.17}$$

Obtaining the global minima (i.e. the smallest root for T) over $\hat{V}$ appears extremely difficult in closed form and has no been achieved. In the case of even two qubits there are 15 parameters to be optimised over.

11.3.2 Generalising the ML Bound

In this section the correspondence between the first and second fundamental problems of quantum control, as described in this chapter, is used to obtain a generalisation of the ML bound.

Given any right invariant Finsler metric $F_{\hat{U}}$ on $SU(n)$ defined by a Mincowki norm F on $\mathfrak{su}(n)$ one can define the corresponding length functional:

$$L[\hat{U}_t] = \int_0^T F_{\hat{U}_t}\left(\frac{d\hat{U}_t}{dt}\right)dt = \int_0^T F\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger\right)dt \tag{11.18}$$

Identically, a right invariant action functional can be defined in terms of any positive homogeneous function $F : \mathfrak{su}(n) \to \mathbb{R}$:

$$L[\hat{U}_t] = \int_0^T F_{\hat{U}_t}\left(\frac{d\hat{U}_t}{dt}\right)dt = \int_0^T F\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger\right)dt \tag{11.19}$$

Theorem 11.3.2. *For any system which satisfies the Schrödinger equation and is constrained such that* $F\left(-i\hat{H}_t\right) = 1$ *the length of any curve is equal to the time the time evolution operator* $\hat{U}_t$ *of such a system takes to traverse the curve:*

$$\begin{aligned} L[\hat{U}_t] &= \int_0^T F_{\hat{U}_t}\left(\frac{d\hat{U}_t}{dt}\right) dt \\ &= \int_0^T F\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger\right) dt \\ &= \int_0^T F\left(-i\hat{H}_t\right) dt = T \end{aligned} \tag{11.20}$$

In order to obtain generalisation of the Margolis-Levitin bound one can suppose a gate $\hat{O}$ is implemented in a quantum system with time independent Hamiltonian $\hat{K}$, and that the gate takes time T to implement. That is:

$$\hat{U}_T = e^{-iT\hat{K}} = \hat{O} \tag{11.21}$$

This implies (by taking logs and rearranging):

$$\hat{K} = \frac{i}{T}\log(\hat{O}) \tag{11.22}$$

It is clear that setting the Hamiltonian $\hat{K} = \frac{i}{T}\log(\hat{O})$ implements $\hat{O}$ at time $t = T$. This is because $\hat{U}_t = e^{\frac{t}{T}\log(\hat{O})}$ which at time $t = T$ comes to $e^{\log(\hat{O})} = \hat{O}$. We can now find the action of the curve $\hat{U}_t$ connecting $\hat{I}$ to $\hat{O}$ along a time independent trajectory as follows. Let F be a right invariant PH function on $T\,SU(N)$. Then:

$$\begin{aligned} S[\hat{U}_t] &= \int_0^T F_{e^{-it\hat{K}}}\left(\frac{d}{dt}e^{-it\hat{K}}\right) dt \\ &= \int_0^T F\left(-i\hat{K}\right) dt \\ &= TF\left(-i\frac{i}{T}\log(\hat{O})\right) \\ &= F\left(\log(\hat{O})\right) \end{aligned} \tag{11.23}$$

In the final step the Ts cancel due to the assumed homogeneity of F. As any such action is invariant under reparameterisation of the one parameter subgroup connecting $\hat{I}$ to $\hat{O}$, we have the following theorem:

Theorem 11.3.3. *Given any PH function* $F : \mathfrak{su}(N)$*, then any time independent, finite dimensional quantum system with Hamiltonian* $\hat{H}$ *such that* $\hat{U}_T = \hat{O}$ *satisfies:*

$$T = \frac{F\left(\log(\hat{O})\right)}{F\left(-i\hat{H}\right)} \tag{11.24}$$

Proof. Any two parameterisations of any curve must yield the same value for the action, as F is a PH function. $TF\left(-i\hat{H}\right)$ and $F\left(\log(\hat{O})\right)$ are two different formulae for the action for two parametrisation of the same curve, hence they must be equal. □

It should be noted that an implicit assumption about F has been made. It is assumed that F is non-singular at $\log(\hat{O})$ and non-zero at $\log(-i\hat{H})$ so that the expression for T is finite. Theorem 11.3.3 has a corollary.

Theorem 11.3.4. *If the Hamiltonian is constrained such that $F(-i\hat{H}) = \kappa$ for some $\kappa \in \mathbb{R}$ then:*

$$T = \frac{1}{\kappa} F\left(\log(\hat{O})\right) \tag{11.25}$$

Motivation For Proof Technique

At this point we note that 11.25 could be obtained more simply by taking logs and applying F to both sides of:

$$e^{-iT\hat{H}} = \hat{O} \tag{11.26}$$

however, as the result 11.3.2 applies to time dependent systems, seeking a proof using a length functional puts both types of system on the same footing and provides and example of assessing the time to traverse a specific trajectory. The fact that a simpler, alternative method is available in the case of time independent trajectories would not provide an example of how arbitrary trajectories could be assessed.

Obtaining the Bound

Now one seeks the time required to implement the orthogonalising gate $\hat{T}$ in a system constrained such that $F\left(-i\hat{H}\right) = \kappa$. Only a two level system will be considered here the proof in the n-level case is almost identical but slightly less clear. By 11.25 this time is:

$$\begin{aligned} T &= \frac{1}{\kappa} F\left(\log(\hat{T})\right) \\ &= \frac{1}{\kappa} F\left(\log(\hat{V} O \hat{V}^\dagger)\right) \\ &= \frac{1}{\kappa} F\left(\hat{V} \log(O) \hat{V}^\dagger\right) \end{aligned} \tag{11.27}$$

Let $G^{(|\psi\rangle)} : \mathfrak{su}(N) \to \mathbb{R}$ be defined by:

$$G^{(|\psi\rangle)}(-i\hat{H}) = \frac{\langle\psi|\hat{H} - E_0\hat{I}|\psi\rangle}{\langle\psi|\psi\rangle} = \overline{E} - E_0 \tag{11.28}$$

where E_0 is the lowest eigenvalue of $\hat{H}$. This function is a special case ($p = 1$) of a more general $G_p, p > 0$:

$$G_p^{(|\psi\rangle)}(-i\hat{H}) = \frac{\left(\langle\psi|(\hat{H} - E_0\hat{I})^p|\psi\rangle\right)^{1/p}}{\langle\psi|\psi\rangle} \tag{11.29}$$

G_1 is the energy expectation and G_2 is the energy uncertainty. We have that $\forall p > 0$, $\forall \lambda > 0$:

$$G_p^{(|\psi\rangle)}(\lambda(-i\hat{H})) = \frac{(\langle\psi|(\lambda\hat{H} - \lambda E_0\hat{I})^p|\psi\rangle)^{1/p}}{\langle\psi|\psi\rangle} \tag{11.30}$$
$$= \frac{(\lambda^p\langle\psi|(\hat{H} - E_0\hat{I})^p|\psi\rangle)^g 1/p}{\langle\psi|\psi\rangle}$$
$$= \lambda G_p^{(|\psi\rangle)}(-i\hat{H})$$

Hence all the G_p are positive homogeneous functions on $\mathfrak{su}(N)$ for a fixed state and value of p, so 11.3.3 can be applied. Furthermore, as G_p each depend only on the spectrum of their argument, they are invariant under unitary conjugation. Applying these facts to 11.20 yields the result:

$$T = \frac{1}{\kappa} G_p^{(|\psi\rangle)}\left(\hat{V}\log(\hat{O})\hat{V}^\dagger\right) \tag{11.31}$$
$$= \frac{1}{\kappa} G_p^{(|\psi\rangle)}\left(\log(\hat{O})\right)$$

One further finds:

$$\log\left(\hat{O}\right) = \frac{\pi}{2}\begin{pmatrix} 0 & ie^{-i\theta} \\ ie^{i\theta} & 0 \end{pmatrix} \tag{11.32}$$

thus:

$$\begin{aligned}
T &= \frac{1}{\kappa} G_p^{(\hat{V}|0\rangle)}\left(\hat{V}\log(\hat{O}\hat{V}^\dagger)\right) \\
&= \frac{1}{\kappa} G_p^{(\hat{V}|0\rangle)}\left(\frac{\pi}{2}\hat{V}\begin{pmatrix} 0 & ie^{-i\theta} \\ ie^{i\theta} & 0 \end{pmatrix}\hat{V}^\dagger\right) \\
&= \frac{\pi}{2\kappa} G_p^{(\hat{V}|0\rangle)}\left(\hat{V}\begin{pmatrix} 0 & ie^{-i\theta} \\ ie^{i\theta} & 0 \end{pmatrix}\hat{V}^\dagger\right) \\
&= \frac{\pi}{2\kappa}\left(\langle 0|\hat{V}^\dagger\left(i\hat{V}\left(\begin{pmatrix} 0 & ie^{-i\theta} \\ ie^{i\theta} & 0 \end{pmatrix}\right)\hat{V}^\dagger + \hat{I}\right)^p \hat{V}|0\rangle\right)^{1/p} \\
&= \frac{\pi}{2\kappa}\left(\langle 0|\begin{pmatrix} 1 & -e^{-i\theta} \\ -e^{i\theta} & 1 \end{pmatrix}^p |0\rangle\right)^{1/p} \\
&= \frac{\pi}{2\kappa}\left((10\cdots 0)\begin{pmatrix} 2^{p-1} & -2^{p-1}e^{-i\theta} \\ -2^{p-1}e^{i\theta} & 2^{p-1} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\right)^{1/p} \\
&= \frac{\pi}{\kappa 2^{1/p}}
\end{aligned}$$

which in the $p = 1$ case yields the ML bound:

$$T = \frac{\pi}{2(\bar{E} - E_0)} \tag{11.33}$$

noting the $\hbar$ is only absent due to non-dimensionalisation. This formula is in general equal to the bound given in both [68, 70]. One notes that this does not yield the Mandelshtam-Tamm inequality relation in the $p = 2$ case.

Mandelshtam-Tamm inequality

One can also define:

$$L^{(|\psi\rangle)}\left(-i\hat{H}\right) = \frac{(\langle\psi|(\hat{H}-\bar{E}\hat{I})^2|\psi\rangle)^{1/2}}{\langle\psi|\psi\rangle} \tag{11.34}$$

where $\bar{E} = \langle\psi|\hat{H}|\psi\rangle$. One can confirm that this is a positive homogeneous function on $\mathfrak{su}(n)$ by an almost identical argument to the ML bound case which is omitted. This function is equal to the energy uncertainty in the state $|\psi\rangle$.

Now the Mandelshtam-Tamm inequality can be proven in an almost identical fashion to the ML bound above by applying 11.25:

$$\begin{aligned} T &= \frac{1}{\kappa} L^{\hat{V}|0\rangle}\left(\hat{V}\log(\hat{O})\hat{V}^\dagger\right) \\ &= \frac{1}{\kappa}(\langle 0|(i\log(\hat{O}) - \bar{E}\hat{I})^2|0\rangle)^{1/2} \\ &= \frac{1}{\kappa}(\langle 0|(i\log(\hat{O}) - i\langle 0|\log(\hat{O})|0\rangle\hat{I})^2|0\rangle)^{1/2} \\ &= \frac{1}{\kappa}(\langle 0|(i\log(\hat{O}))^2|0\rangle)^{1/2} \\ &= \frac{\pi}{2\kappa}\left(\langle 0|\begin{pmatrix} 0 & e^{-i\theta} \\ e^{i\theta} & 0\end{pmatrix}^2|0\rangle\right)^{1/2} \\ &= \frac{\pi}{2\kappa}\left(\langle 0|\hat{I}|0\rangle\right)^{1/2} \\ &= \frac{\pi}{2\kappa} \\ &= \frac{\pi}{2\Delta E} \end{aligned} \tag{11.35}$$

which proves the Mandelshtam-Tamm inequality.

The Operator Norm

The method can be used to obtain a large family of novel bounds. The operator norm $||\cdot||_{op}$ of a matrix is defined by

$$||\hat{A}||_{op}^2 = \max\left\{\frac{\langle\psi|\hat{A}^\dagger\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle}, \forall|\psi\rangle\in\mathbb{C}^N\right\} \tag{11.36}$$

This is equal to the largest singular value of $\hat{A}$, often written $\sigma_{max}(\hat{A})$. It is unitarily invariant, that is $||\hat{V}A\hat{V}^\dagger||_{op} = ||\hat{A}||_{op}$ for any unitary $\hat{V}$ which can be easily confirmed as it depends only on the spectrum of its argument.

By applying an identical argument to the type above in 11.3.2, 11.3.2 one finds:

$$\begin{aligned}
T &= \frac{1}{\kappa} G_{op}(\log(\hat{O}_{\hat{V}})) \qquad (11.37) \\
&= \frac{1}{\kappa} \left|\left| \frac{\pi}{2}(i\hat{V}\hat{O}\hat{V}^{\dagger} + \hat{I}) \right|\right|_{op} \\
&= \frac{1}{\kappa} \left|\left| \frac{\pi}{2}(i\hat{O} + \hat{I}) \right|\right|_{op} \\
&= \frac{\pi}{2\kappa} \left|\left| i\hat{O} + \hat{I} \right|\right|_{op} \\
&= \frac{\pi}{2\kappa} \sigma_{max}(i\hat{O} + \hat{I}) = \\
&= \frac{\pi}{\kappa} \\
&= \frac{\pi}{E_{\max} - E_0}
\end{aligned}$$

(11.38)

11.3.3 Pros and Cons of Method

This method for obtaining speed limits for reaching an orthogonal state sets many existing known bounds on an equal footing by showing that they are all special cases of a far more general statement. It further vindicates the intuition expressed in 5.19. However, it is significantly more involved to apply, at lease conceptually, that the existing proof of the Margolis Levitin theorem. Furthermore, it is unfortunately also not possible, in the current formulation, to apply proof technique to infinite dimensional systems where the existing ML bound applies unchanged.

Chapter 12

Further Work, Outlook and Conclusions

12.1 Efficacy of Zermelo Navigation Based Methods For Quantum Control and Analysis of the QSL

In this thesis a method is presented for obtaining a differential equation, an EP equation in $\mathfrak{su}(n)$, satisfied by the optimal Hamiltonian implementing a desired quantum gate. This equation has also been shown to be solvable in the isotropic case, when h is a multiple of the Killing form. This case is of particular significance as this it is physically plausible that all control fields will face equal restrictions. However, it was already possible to obtain optimal Hamiltonians for quantum gates in Constrained systems using the methods presented in the work of Carlini [4]. However, it might be argued that the methods presented here are conceptually simpler and geometrical as the optimal trajectories have been revealed to be sub-Finsler geodesics.

The result 6.20 determining the geodesics and optimal Hamiltonian in the case where h is a multiple of the Killing form is not dependent on the dimension of the system to which it applies. This feature is powerful as it allows for the analysis of systems of arbitrary number of qubits to be performed. However, in order to obtain practical optimal Hamiltonians in systems with additional constraints one must include a very large number of Lagrange multipliers. In the example of an n-spin Heisenberg spin chain with one control field per site, the number of Lagrange multipliers is $(\dim(\mathfrak{su}(2^n)) - n) = (2^{2n} - 1 - n)$ which clearly grows at order $O(2^{2n})$.

Studying methods for obtaining numerical solutions to the time optimality equations of Heisenberg spin chains will require further work. It is the hope of the author that such an investigation will clarify the significance of the number of Lagrange multipliers and determine if the rate of there growth is significant obstacle.

One of the other main advantages of the methods in this work in the ability to obtain the optimal time for $\hat{U}_t$ to traverse arbitrary trajectories on $SU(n)$ in constrained quantum system. This contribution is an entirely new capacity in quantum control; the author know of no other general method achieving this. Using the metric 6.5, one can find the length of arbitrary curves on $SU(n)$, and thus the optimal times for $\hat{U}_t$ to traverse that curve. This is practically relevant as not all real quantum systems can traverse all possible trajectories on $SU(n)$, even though the system might be controllable. Controllability is only a statement about which points on $SU(n)$ can be reached in some finite time, not about which trajectories can be traversed.

Fully exploring the application of Shen's theorem and Zermelo navigation more generally in scenarios in which not every trajectory can be achieved has not been achieved in this work. The only example 6.5 has been the analysis is that of the optimal times for traversing one parameter subgroups of $SU(n)$ which corresponds to the problem of finding optimal, time independent, controls culminating in 6.56. Further work could analyse systems of the form:

$$\frac{d\hat{U}_t}{dt} = -i\left(\hat{H}_0 + \sum_{k=0}^{N} f_k(\vec{a},t)\hat{H}_k\right)\hat{U}_t \tag{12.1}$$

where f_k is a parametrised family of functions of times, parametrised by the components of the vector $\vec{a}$. For example, it is not possible to control exactly the output of a laser in reality. It is only really possible to tune some aspects of the output, for example the polarisation and amplitude; and even these cannot be varied arbitrarily quickly. The effect of a laser on a single spin, or simple molecule can be very simplistically modelled by the following term in the system's Hamiltonian:

$$A\sin(\omega t)\sigma_z + B\cos(\omega t)\sigma_y \tag{12.2}$$

and in such an example $\vec{v} = \begin{pmatrix} A \\ B \\ \omega \end{pmatrix}$. The overall Hamiltonian could, for example, be:

$$\hat{H}_t = \kappa\sigma_x + A\sin(\omega t)\sigma_z + B\cos(\omega t)\sigma_y \tag{12.3}$$

Here only a finite (3 in this case) number of parameters require to be chosen. If a condition can be found on $\vec{v}$ so that a desired gate is implemented, a method is still required for finding the time optimal values for $\vec{v}$. By applying the results of 6 one easily sees that this condition is:

$$\frac{d}{dv_k}\int_0^T \frac{1}{2}F\left(-i\hat{H}_t\right)^2 dt = 0 \tag{12.4}$$

More precisely, this condition would produce stationary values of $\vec{v}$, a second order derivative condition would be needed to confirm that a specific stationary value of $\vec{v}$ was globally optimal.

12.2 Potential For Application In Natural Computation

This section is primarily a discussion of the potential applications of Zermelo navigation in other areas of natural computation. The possibility of geometric control as a unifying basis for some areas of Natural computation are discussed as are the limitations of this concept.

One can consider more theories of computation than the classical Turing theory and theories based on the non-relativistic, finite dimensional quantum mechanics of pure states but on other aspects of physics. There is much interest in putting the theory of computation on physical grounds [8, 9, 10, 11] and various perspectives have been taken.

Many, but not all, systems which are studied in natural computing have a common mathematical foundation. That common foundation is dynamical systems, see [119] for a conceptual discussion of computations embedded in many types of dynamical systems. See [120, 121, 122] for more specific details of computations embedded into specific dynamical systems. One class of dynamical systems with particular relevance in physics are those where time is continuous and the state space is a topological space or manifold, examples abound. By no means the least important example is that of the geometric quantum mechanics description of quantum systems studied in this thesis.

In the Turing theory of computation the 'speed' of a computation is typically equated with computational time complexity. However in theories of computation based on physical foundations, it is more appropriate to seek physical times in order to quantify the speed of operation of a computer.

It has been shown that there are more examples of physical systems which are considered as information processing systems to which geometric control can be applied. Examples include:

- Classical mechanical systems with symplectic manifolds as their state spaces and Hamiltonian time evolution
- Probabilistic systems such as finite state, continuous time Markov chains, for which the state space is a probability simplex (which is a manifold) and the Master equation as time evolution

However, these examples have not be expounded in detail.

While geometric control does not apply to many of the dynamical systems in use as models of computation, it does apply to a broad range of them. It is the belief of the author that the type of system to which it applies indicates that it could form the basis of a fully developed theory of time complexity in physics based models of computation if not non-standard/natural computation more broadly.

Computation Embodied In a Dynamical System

In the quantum systems studied in this thesis a meaning is taken for the idea that a system implements a computation. This prescription is that $\hat{U}_t = \hat{O}$ where $\hat{O}$ is some desired time evolution. However, it is not immediately clear how this concept should be generalised to other types of system in natural computation in order to facilitate the comparison of the speed and power of different physics based models of computation. Such a prescription is required if one is to answer the question "what is the least time a system can implement a computation" as one must first have a way to say that it implements it at all.

12.2.1 Systems To Which Zermelo Navigation Does Not Apply

There are many systems in computing which are not amenable to analysis using Zermelo navigation or even geometric control of any kind known to the author. These are the systems which have discrete time evolution such as:

- Discrete time Markov chains [123]
- Quantum Cellular Automata [124]

and ones for which the state space (or both state and time) is discrete:

- Binary Cellular Automata [125]
- Finite State Automata [126]

12.2.2 Time Complexity and Computability In Physical Models of Computation

There is an analogy between controlability and computability and an analogy between time optimal control and complexity.

In computer science, very broadly speaking, one speaks of the speed of classical computers in terms of time complexity. One speaks of which overall time evolutions a certain model of computation permits as computability. Furthermore, one broadly calls computers which can implement every (in some context dependent sense) overall time evolution as 'universal'. The analogy with control theory is very clear:

Theory Of Computation	Control Theory For NSC
Computability	Reachable Set
Universality	Controlability
Time Complexity	Time Optimal Control

12.3 Conclusion

Zermelo navigation is a powerful tool for quantum control and can produce novel speed limit results for quantum computing and assess properties of quantum systems not previously possible. Ultimately these contributions work towards a more satisfactory answer to the question of the limits to computation posed in [66] and elsewhere.

Zermelo navigation methods only apply to affine control systems on a manifold. This includes many physical systems, but not to all systems considered in unconventional/natural computation. However, it does apply to many 'physics like' models of computation as these are frequently dynamical systems with a smooth manifold of states and continuous time evolution described by a differential equation. Such systems include classical mechanical systems with symplectic manifolds as their phase spaces and probabilistic systems (continuous time Markov chains with finitely many states) for which probability simplexes are the state spaces.

Geometric control theory has a clear analogy with concepts of interest in classical computer science. As such the author believes that it could play a foundational role in assessing physics based models of computation beyond quantum mechanics. The two mains obstacles to this program the author foresees are:

- In relativistic systems time is no longer an independent parameter as it is in the Zermelo navigation problem so the problem would need adapting
- In quantum field theories realistic systems are necessarily infinite dimensional. As such Zermelo navigation in infinite dimensional Lie groups would need to be developed. This would incur significant mathematical complexity.

Bibliography

[1] Benjamin Russell and Susan Stepney. Applications of finsler geometry to speed limits to quantum information processing. *International Journal of Foundations of Computer Science*, 25(04):489–505, 2014.

[2] Benjamin Russell and Susan Stepney. Zermelo navigation and a speed limit to quantum information processing. *Phys. Rev. A*, 90:012303, Jul 2014.

[3] Benjamin Russell and Susan Stepney. Zermelo navigation in the quantum brachistochrone. *Journal of Physics A: Mathematical and Theoretical*, 48(11):115303, 2015.

[4] Alberto Carlini, Akio Hosoya, Tatsuhiko Koike, and Yosuke Okudaira. Time-optimal unitary operations. *Phys. Rev. A*, 75:042308, Apr 2007.

[5] Alberto Carlini, Akio Hosoya, Tatsuhiko Koike, and Yosuke Okudaira. Time-optimal quantum evolution. *Phys. Rev. Lett.*, 96:060503, Feb 2006.

[6] R.P. Feynman, A.J.G. Hey, and R.W. Allen. *Feynman Lectures on Computation*. Perseus Books, 2000.

[7] R.P. Feynman. There's plenty of room at the bottom. *Microelectromechanical Systems, Journal of*, 1(1):60–66, March 1992.

[8] Tommaso Toffoli. Physics and computation. *International Journal of Theoretical Physics*, 21(3-4):165–175, 1982.

[9] M. Mézard and A. Montanari. *Information, Physics, and Computation*. OUP Oxford, 2009.

[10] Subhash Kak. Information, physics, and computation. *Foundations of Physics*, 26(1):127–137, 1996.

[11] Norman Margolus. Physics-like models of computation. *Physica D: Nonlinear Phenomena*, 10(12):81 – 95, 1984.

[12] M.A. Nielsen and I.L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[13] V. Jurdjevic. *Geometric Control Theory*. Cambridge University Press, 1997.

[14] Ruth F Curtain and Anthony J Pritchard. *Infinite dimensional linear systems theory*, volume 8. Springer-Verlag, 1978.

[15] H.J. Zwart. *Geometric Theory for Infinite Dimensional Systems.* Lecture Notes in Control and Information Sciences. Springer-Verlag, 1989.

[16] Harris F. Goldstein and Carl M. Bender. Relativistic brachistochrone. *Journal of Mathematical Physics*, 27(2), 1986.

[17] Fabio Giannoni and Paolo Piccione. The arrival time brachistochrones in general relativity. *The Journal of Geometric Analysis*, 12(3):375–423, 2002.

[18] J. Gemmer, R. Umble, and M. Nolan. Generalizations of the Brachistochrone Problem. *ArXiv*, December 2006.

[19] D. E. Bruschi, C. Sabín, P. Kok, G. Johansson, P. Delsing, and I. Fuentes. Towards universal quantum computation through relativistic motion. *ArXiv*, November 2013.

[20] S Pakvasa, W Simmons, and X Tata. Relativistic quantum computing. *arXiv*, 1999.

[21] D. Bump. *Lie Groups.* Springer, 2004.

[22] K.-Y. Lee and H. F. Chau. Relation between quantum speed limits and metrics on U(n). *Journal of Physics A Mathematical General*, 46(1):015305, January 2013.

[23] X. Cheng and Z. Shen. *Finsler Geometry: An Approach via Randers Spaces.* Springer, 2013.

[24] Gunnar Randers. On an asymmetrical metric in the four-space of general relativity. *Phys. Rev.*, 59:195–199, Jan 1941.

[25] G.W. Gibbons and C.M. Warnick. The geometry of sound rays in a wind. *Contemporary Physics*, 52(3):197–209, 2011.

[26] P.L. Antonelli and R. Miron. *Lagrange and Finsler Geometry: Applications to Physics and Biology.* Fundamental Theories of Physics. Springer, 1995.

[27] R. Ingarden. *On the Geometrically Absolute Optical Representation in the Electron Microscope.* Prace Wrocławskiego Towarzystwa Naukowego. Państwowe Wydawn. Naukowe, 1957.

[28] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian Geometry.* Universitext (1979). U.S. Government Printing Office, 2004.

[29] Z. Shen. Finsler Metrics with K=0 and S=0. *Canadian J. of Math.*, 55, 2003.

[30] E. Zermelo. über das navigationsproblem bei ruhender oder veränderlicher windverteilung. *Z. Angew. Math. Mech.*, 11:114–124, 1931.

[31] C. Carathéodory. *Calculus of Variations and Partial Differential Equations of the First Order*. Chelsea Publishing Series. American Mathematical Society, 1999.

[32] P.K. Jain, K. Ahmad, and O.P. Ahuja. *Functional Analysis*. New Age International, 1995.

[33] David Bao, Colleen Robles, and Zhongmin Shen. Zermelo navigation on riemannian manifolds. *Journal of Differential Geometry*, 66(3):377–435, 03 2004.

[34] Dorje C. Brody and Lane P. Hughston. Geometric quantum mechanics. *Journal of Geometry and Physics*, 38(1):19 – 53, 2001.

[35] John Von Neumann. *Mathematical Foundations of Quantum Mechanics*. Investigations in physics. Princeton University Press, 1955.

[36] P.A.M. Dirac. *Lectures on Quantum Mechanics*. Dover Publications, 2013.

[37] A. W. Knapp. *Lie Groups Beyond an Introduction*. Progress in Mathematics. Birkhäuser Boston, 2002.

[38] J.E. Humphreys. *Linear Algebraic Groups*. Springer, 1975.

[39] B. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Springer, 2003.

[40] Roderich Tumulka. Povm (positive operator value measure). In Daniel Greenberger, Klaus Hentschel, and Friedel Weinert, editors, *Compendium of Quantum Physics*, pages 480–484. Springer, 2009.

[41] V.B. Braginsky, V.B. Braginskiĭ, F.Y. Khalili, and K.S. Thorne. *Quantum Measurement*. Cambridge University Press, 1995.

[42] J.M. Lee. *Introduction to Topological Manifolds*. Springer, 2000.

[43] A. Arvanitogeōrgos. *An Introduction to Lie Groups and the Geometry of Homogeneous Spaces*. Student mathematical library. American Mathematical Soc., 2003.

[44] B. Li, Z. Yu, S. Fei, and X. Li-Jost. Time optimal quantum control of two-qubit systems. *Science China Physics, Mechanics, and Astronomy*, 56:2116–2121, November 2013.

[45] T. Schulte-Herbrüggen, A. Spörl, N. Khaneja, and S. J. Glaser. Optimal control for generating quantum gates in open dissipative systems. *Journal of Physics B Atomic Molecular Physics*, 44(15):154013, August 2011.

[46] Navin Khaneja, Steffen J. Glaser, and Roger Brockett. Sub-riemannian geometry and time optimal control of three spin systems: Quantum gates and coherence transfer. *Phys. Rev. A*, 65, Jan 2002.

[47] Benjamin Russell and Susan Stepney. Geometric methods for analysing quantum speed limits: time-dependent controlled quantum systems with constrained control functions. In *Unconventional Computation and Natural Computation 2013, Milan, Italy, July, 2013*, LNCS. Springer, 2013.

[48] A.A. Agrachev and Y. Sachkov. *Control Theory from the Geometric Viewpoint*. Control theory and optimization. Springer, 2004.

[49] A.F.J. Levi. *Applied Quantum Mechanics*. Cambridge University Press, 2006.

[50] P. E. G. Assis and A. Fring. The quantum brachistochrone problem for non-Hermitian Hamiltonians. *Journal of Physics A Mathematical General*, 41(24):244002, June 2008.

[51] D. D'Alessandro. *Introduction to Quantum Control and Dynamics*. Taylor & Francis, 2007.

[52] S. Blanes, F. Casas, J.A. Oteo, and J. Ros. The Magnus expansion and some of its applications. *Physics Reports*, 470(5-6):151–238, 2009.

[53] S Blanes, F Casas, J A Oteo, and J Ros. Magnus and Fer expansions for matrix differential equations: the convergence problem. *Journal of Physics A: Mathematical and General*, 31(1):259, 1998.

[54] The brachistochrone. In *Essentials of Mathematica*, pages 285–287. Springer New York, 2007.

[55] X. Wang, M. Allegra, K. Jacobs, S. Lloyd, C. Lupo, and M. Mohseni. Quantum brachistochrone curves as geodesics: obtaining accurate control protocols for time-optimal quantum gates. *ArXiv*, August 2014.

[56] Susan Stepney, Samuel L. Braunstein, John A. Clark, Andy Tyrrell, Andrew Adamatzky, Robert E. Smith, Tom Addis, Colin Johnson, Jonathan Timmis, Peter Welch, Robin Milner, and Derek Partridge. Journeys in non-classical computation i: A grand challenge for computing research. *International Journal of Parallel, Emergent and Distributed Systems*, 20(1):5–19, 2005.

[57] Susan Stepney, Samuel L. Braunstein, John A. Clark, Andy Tyrrell, Andrew Adamatzky, Robert E. Smith, Tom Addis, Colin Johnson, Jonathan Timmis, Peter Welch, Robin Milner, and Derek Partridge. Journeys in non-classical computation ii: initial journeys and waypoints. *International Journal of Parallel, Emergent and Distributed Systems*, 21(2):97–125, 2006.

[58] D. Bouwmeester, A.K. Ekert, and A. Zeilinger. *The Physics of Quantum Information: Quantum Cryptography, Quantum Teleportation, Quantum Computation*. Springer, 2000.

[59] Clare Horsman, Susan Stepney, Rob C. Wagner, and Viv Kendon. When does a physical system compute? *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 470(2169), 2014.

[60] J. C. Baez and M. Stay. Physics, Topology, Logic and Computation: A Rosetta Stone. *ArXiv*, March 2009.

[61] James Ladyman. What does it mean to say that a physical system implements a computation? *Theoretical Computer Science*, 410(45):376 – 383, 2009. Computational Paradigms from Nature.

[62] Elliott H. Lieb and Derek W. Robinson. The finite group velocity of quantum spin systems. *Communications in Mathematical Physics*, 28(3):251–257, 1972.

[63] R. Sims and D. Ueltschi. *Entropy and the Quantum: Arizona School of Analysis with Applications, March 16-20, 2009, University of Arizona*. American Mathematical Society, 2010.

[64] Norman Margolus and Lev B. Levitin. The maximum speed of dynamical evolution. *Physica*, D120:188–195, 1998.

[65] H.M. Wiseman and G.J. Milburn. *Quantum Measurement and Control*. Cambridge University Press, 2010.

[66] Seth Lloyd. Ultimate physical limits to computation. *Nature*, 406(6799):1047–1054, 1999.

[67] M. Andrecut and M. K. Ali. The adiabatic analogue of the Margolus–Levitin theorem. *Journal of Physics A: Mathematical and General*, 37(15):L157, 2004.

[68] B. Zieliński and M. Zych. Generalization of the Margolus-Levitin bound. 74(3):034301, September 2006.

[69] L. B. Levitin and T. Toffoli. Fundamental Limit on the Rate of Quantum Dynamics: The Unified Bound Is Tight. *Physical Review Letters*, 103(16):160502, October 2009.

[70] Q. Cao, Y.-X. Chen, and J.-L. Li. Covariant versions of Margolus-Levitin Theorem. *ArXiv*, May 2008.

[71] P. Busch. The Time-Energy Uncertainty Relation. In J. G. Muga, R. Sala Mayato, and I. L. Egusquiza, editors, *Time in Quantum Mechanics*, page 69, 2002.

[72] Y. Aharonov and D. Bohm. Time in the quantum theory and the uncertainty relation for time and energy. *Phys. Rev.*, 122:1649–1658, Jun 1961.

[73] R.O.N. Wells. *Differential Analysis on Complex Manifolds*. Springer Verlag, 1980.

[74] P. J. Jones and P. Kok. Geometric derivation of the quantum speed limit. *ArXiv*, 82(2):022107, August 2010.

[75] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry*. Wiley, 1996.

[76] Florian Frwis, Gebhard Grbl, and Markus Penz. Fleming's bound for the decay of mixed states. *Journal of Physics A: Mathematical and Theoretical*, 41(40):405201, 2008.

[77] D. C. Brody. Elementary derivation for passage times. *Journal of Physics A Mathematical General*, 36:5587–5593, May 2003.

[78] M. Zwierz. Comment on "Geometric derivation of the quantum speed limit". *Physical Review A*, 86:016101, 2012.

[79] M. M. Taddei, B. M. Escher, L. Davidovich, and R. L. de Matos Filho. Quantum Speed Limit for Physical Processes. *Physical Review Letters*, 110(5):050402, February 2013.

[80] S. Deffner and E. Lutz. Quantum Speed Limit for Non-Markovian Dynamics. *Physical Review Letters*, 111(1):010402, July 2013.

[81] A. Garon, S. J. Glaser, and D. Sugny. Time-optimal control of su(2) quantum operations. *Phys. Rev. A*, 88:043422, Oct 2013.

[82] Carl M. Bender, Dorje C. Brody, Hugh F. Jones, and Bernhard K. Meister. Faster than Hermitian quantum mechanics. *Phys.Rev.Lett.*, 98:040403, 2007.

[83] T. Caneva, M. Murphy, T. Calarco, R. Fazio, S. Montangero, V. Giovannetti, and G. E. Santoro. Optimal control at the quantum speed limit. *Phys. Rev. Lett.*, 103:240501, Dec 2009.

[84] Jacob Bekenstein. Universal upper bound on the entropy-to-energy ratio for bounded systems. *Phys. Rev. D*, 23:287–298, Jan 1981.

[85] JacobD. Bekenstein. How does the entropy/information bound work? *Foundations of Physics*, 35(11):1805–1823, 2005.

[86] Jacob D Bekenstein. Black holes and information theory. *Contemporary Physics*, 45(1):31–43, 2004.

[87] R. Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3):183–191, July 1961.

[88] S. Lloyd. Quantum limits to the measurement of spacetime geometry. *arXiv*, May 2005.

[89] Q. Cao, Y.-X. Chen, and J.-L. Li. Covariant versions of Margolus-Levitin Theorem. *ArXiv*, May 2008.

[90] M. A. Nielsen. A geometric approach to quantum circuit lower bounds. *Quant. Info. Comp*, 6:213–262, 2006.

[91] H. Cendra, J.E. Marsden, and T.S. Ratiu. *Lagrangian Reduction by Stages*. Number v. 722. American Mathematical Soc., 2001.

[92] Darryl D Holm, Jerrold E Marsden, and Tudor S Ratiu. The Euler-Poincaré Equations and Semidirect Products with Applications to Continuum Theories. *Advances in Mathematics*, 137(1):1 – 81, 1998.

[93] R. Berndt. *Representations of Linear Groups: An Introduction Based on Examples from Physics and Number Theory*. Vieweg, 2007.

[94] Aleksandr Aleksandrovich Kirillov. *Lectures on the Orbit Method*. American Mathematical Society, 2004.

[95] Wang-Sang Koon and Jerrold E. Marsden. Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction, 1996.

[96] W. Pfeifer. *The Lie Algebras Su(n): An Introduction*. Birkhäuser Basel, 2003.

[97] W. Fulton and J. Harris. *Representation Theory: A First Course*. Springer New York, 1991.

[98] Dorje Brody and David Meier. Solution to the quantum Zermelo navigation problem. *Physical Review Letters*, 114, 2015.

[99] Colleen Robles. Geodesics in randers spaces of constant curvature.

[100] Libing Huang and Xiaohuan Mo. On geodesics of finsler metrics via navigation problem. *Proc. Amer. Math. Soc. 139 (2011), 3015-3024*, 2011.

[101] J. Stillwell. *Naive Lie Theory*. Springer, 2008.

[102] Ryozo Yoshikawa and SorinV. Sabau. Kropina metrics and Zermelo navigation on Riemannian manifolds. *Geometriae Dedicata*, 171(1):119–148, 2014.

[103] M. H. Stone. On one-parameter unitary groups in Hilbert space. *Ann. Math*, 33(2):643–648, 1932.

[104] T. A. Loring. Computing a logarithm of a unitary matrix with general spectrum. *ArXiv*, March 2012.

[105] D. J. H.. Garling. *Clifford Algebras: An Introduction.* Cambridge University Press, 2011.

[106] L. Šamaj and Z. Bajnok. *Introduction to the Statistical Physics of Integrable Many-body Systems.* Cambridge University Press, 2013.

[107] Parastoo Habibi, Dariush Latifi, and Megerdich Toomanian. Homogeneous geodesics and the critical points of the restricted Finsler function. *Journal of Contemporary Mathematical Analysis*, 46(1):12–16, 2011.

[108] D. Latifi. Homogeneous geodesics of left invariant Finsler metrics. *ArXiv*, November 2007.

[109] Martin Bauer, Martins Bruveris, and PeterW. Michor. Overview of the geometries of shape spaces and diffeomorphism groups. *Journal of Mathematical Imaging and Vision*, 50(1-2):60–97, 2014.

[110] Enrico Le Donne. Lecture notes on sub-Riemannian geometry. https://sites.google.com/site/enricoledonne/.

[111] G. Stefani. *Geometric Control Theory and Sub-Riemannian Geometry.* 2014.

[112] C. López and E. Martínez. Sub-finslerian metric associated to an optimal control system. *SIAM J. Control Optim.*, 39(3):798–811, September 2000.

[113] N. Khaneja, S. J. Glaser, and R. Brockett. Sub-Riemannian geometry and time optimal control of three spin systems: Quantum gates and coherence transfer. *PRA*, 65(3):032301, March 2002.

[114] Velimir Jurdjevic and Hctor J Sussmann. Control systems on Lie groups. *Journal of Differential Equations*, 12(2):313 – 329, 1972.

[115] D. C. Brody, G. W. Gibbons, and D. M. Meier. Time-optimal navigation through quantum wind. *Arxiv*, October 2014.

[116] Dorje C Brody and David M Meier. Elementary solution to the time-independent quantum navigation problem. *Journal of Physics A: Mathematical and Theoretical*, 48(5):055302, 2015.

[117] Nicoleta Aldea. Zermelo problem of navigation on Hermitian manifolds, 2001.

[118] Shaoqiang Deng. Homogeneous finsler spaces. In *Homogeneous Finsler Spaces*, pages 79–104. Springer New York, 2012.

[119] Susan Stepney. Nonclassical computation: a dynamical systems perspective. chapter 59, pages 1979–2025.

[120] E. Doedel, G. Domokos, I.G. Kevrekidis, and J. Von Neumann. *Modeling and Computations in Dynamical Systems: In Commemoration of the 100th Anniversary of the Birth of John Von Neumann.* World Scientific, 2006.

[121] Jrme Durand-Lose. Computing inside the billiard ball model. In Andrew Adamatzky, editor, *Collision-Based Computing*, pages 135–160. Springer, 2002.

[122] Marston Morse and Gustav A. Hedlund. Symbolic dynamics. *American Journal of Mathematics*, 60(4):pp. 815–866, 1938.

[123] Giuseppe Modica and Laura Poggiolini. *Discrete Time Markov Chains*, pages 168–240. John Wiley and Sons, Ltd, 2012.

[124] C S Lent, P D Tougaw, W Porod, and G H Bernstein. Quantum cellular automata. *Nanotechnology*, 4(1):49, 1993.

[125] T. Toffoli and N. Margolus. *Cellular Automata Machines: A New Environment for Modeling.* 1987.

[126] J.M. Howie. *Automata and languages.* Clarendon Press, 1991.