Optimized Quantum State Transitions:
A Survey of the Quantum Brachistochrone Problem

Robbert W. Scholtens
S2681560
r.w.scholtens@student.rug.nl

University of Groningen
Bachelor's Thesis Physics & Mathematics

July 13th, 2018

Abstract

The quantum brachistochrone problem – to find the time-optimal transition between given initial and final quantum states – is investigated in this bachelor's thesis. First the quantum equivalent of distance (the Fubini-Study metric) is formulated, using geometry of spheres. Together with constraints, it is used to create the quantum action. Functional derivatives are then taken of said action to find equations of motion (eoms). For the unconstrained case (excluding finite energy), these eoms are solved in closed form, whilst general to be solved formulae are obtained for the constrained case. The latter is used to explicitly solve an example quantum brachistochrone problem: a spin-1/2 particle in a controllable magnetic field constrained to an x-y plane. Finally, the link with quantum computing is illustrated through the time-optimization of unitary transformations that the quantum brachistochrone yields us.

Mathematics Supervisor: Prof. Dr. H. Waalkens
Physics Supervisor: Prof. Dr. G. Palasantzas

Contents

1 Introduction

During his tenure as professor of mathematics at the University of Groningen (1694-1705), Johann Bernoulli posed the following problem:

> "Given two points A and B in a vertical plane, what is the curve traced out by a point acted on only by gravity, which starts at A and reaches B in the shortest time."

The problem was subsequently named the *brachistochrone* problem, derived from the Greek words for "shortest time." In true Bernoulli style, he had cracked the problem less than a year later, with the conclusion being that the time-optimal curve is the section of a cycloid.

Although Bernoulli solved it using different means, the problem is commonly posed as an introduction to the optimization of functions using the calculus of variations. In particular, the well-known Euler-Lagrange equation can be used to solve the problem quite easily. And since it is quite an instructive example, we shall briefly go over it below.

We wish to minimize the total travel T. Using that

$$v = \frac{ds}{dt} \quad \Longleftrightarrow \quad dt = \frac{ds}{v}, \tag{1.1}$$

we thus see

$$T = \int_0^T 1\, dt = \int_{s_1}^{s_2} \frac{1}{v}\, ds \overset{*}{=} \int_{x_1}^{x_2} \sqrt{\frac{1 + (y')^2}{2gy}}\, dx, \tag{1.2}$$

where $*$ is the application of $ds^2 = dx^2 + dy^2$, $v = \sqrt{2gy}$ and y is treated as a function of x. It is then a matter of applying the Beltrami identity – a derivative of the Euler-Lagrange equation – which upon solving yields the system

$$\begin{cases} x = A(t - \sin(t)) \\ y = A(1 - \cos(t)) \end{cases} \quad \text{with } A = \frac{1}{4gC^2}, \tag{1.3}$$

where C is an arbitrary constant. This system indeed maps out a cycloid, as predicted by Bernoulli.

Upon shrinking to the quantum level, the similarity shared between the problem outlined above and the *quantum brachistochrone* is that they both seek to minimize a certain transition time. The quantum version, however, seeks to find the shortest possible transition time between two particular states of quantum particle, and the Hamiltonian that goes along with that time. In the paper by Carlini *et al* [4], the problem is formulated thusly:

> "We want to minimize the total amount of time necessary for changing a given initial state $|\psi_i\rangle$ [...] to a given final state $|\psi_f\rangle$, by suitable choice of a (possibly time-dependent) Hamiltonian $H(t)$."

Being the seminal paper on the topic, further research into the quantum brachistochrone (or investigated usages of it) commonly refers back to this paper.

Sadly, the quantum brachistochrone is not as easily solved as the classical one; more mathematical machinery will need to be brought to the table, as well as

equations and interpretations from quantum mechanics, in order to understand the problem and solve it. As this brings with it quite involved derivations, a particular aim of this thesis is to present findings as candidly and completely as possible, using the structure of mathematical writing.

Before we commence with the gist of the thesis, we do a recap of materials the reader should be familiar with. These include the calculus of variations, so that we can take variations of the functions to optimize, and the theory of Lagrange multipliers, so that we can work constraints into our problem.

In the second section of this thesis, we will derive the total action S for a given transition in integral form, with the variable being time. This is the equivalent of finding the total time T in the example above. We will touch on the Fubini-Study line element, a quantum mechanical equivalent of distance, and how we use it to formulate our action. We then discuss constraints to our system, and show how these are implemented using Lagrange multipliers.

Next, we use the aforementioned action to derive the solution of the problem – applying the Euler-Lagrange equation in the classical brachistochrone. This is accomplished by taking the variation of the action functional with respect to the different variables, so that we obtain "equations of motion" for our system. These are then solved for the particular cases of having only one constraint (the so-called "finite energy" constraint), and for arbitrary constraints. The former is solved in closed form, whilst the latter solution depends only on the constraints.

The fourth section will focus on working an example of the constrained version of the quantum brachistochrone: a spin-1/2 particle in a controllable magnetic field. Using the obtained methodology, we solve the example to find the optimal paths and optimal transition time.

Finally, we discuss a particular application of the quantum brachistochrone: quantum computing. Namely, at the core of a quantum computer are qubits which change states (many times) in order to work a calculation. Facilitating the optimal transition time by means of the optimal Hamiltonian thus allows for increasing the speed of quantum computers, as the time spent changing states is minimized.

The goal of this thesis is to provide its reader with a complete picture of the quantum brachistochrone, both mathematically and physically speaking, and as its author I hope that it accomplishes this goal.

Robbert Scholtens, July 2018.

A note on notation

Throughout the report, several notation conventions are utilized. These are outlined here for general reference.

Notation 1.1. We shall use the following abbreviations in this thesis.

1. $\partial_t := d/dt$, $\partial_t^2 := d^2/dt^2$
2. $P := |\psi\rangle\langle\psi|$
3. $\langle A\rangle := \langle\psi|A|\psi\rangle$, where A is any operator
4. $\mathrm{Tr}(A)$ is the *trace* of an operator A, i.e. the sum of its diagonal elements (or the sum of its eigenvalues).
5. $\tilde{H} := H - \mathrm{Tr}(H)/n$
6. $(\Delta E)^2 := \langle H^2\rangle - \langle H\rangle^2$, where the expectations are w.r.t, ψ, as in 3.

Notation 1.2. We take the reduced Planck constant to be unity, i.e. $\hbar \equiv 1$.

Notation 1.3. Whenever an operator is written inside a bra or ket, it is taken to act upon said bra or ket. That is,

$$|A\psi\rangle \equiv A\,|\psi\rangle \quad \text{and} \quad \langle A\psi| \equiv A\,\langle\psi|$$

for all operators A and quantum states ψ.

2 Some required theory

Before we commence with the particular substance of this thesis, it is imperative to discuss some necessary theory. The reasoning for this is twofold. Firstly, the material to be discussed heavily relies on the frameworks we present in this section. As such, without a relatively strong recap of these frameworks, understanding the material might be trickier than it has to be. Secondly, it provides us with an opportunity to present the theory behind the utilized methods. This way, it will become clearer *why* the methods work, and invite application elsewhere as well.

The first subsection is concerned with giving a brief overview of the calculus of variations. That is, the section will introduce the reader to the functional derivative, concretize its connection to the Euler-Lagrange equation and illustrate the chain rule for functional derivatives.

The second subsection treats the method of Lagrangian multipliers. This method gives a very easy way to transform a constrained optimization problem into an unconstrained one, which are much easier to work with in general. Here we will also find a condition for optimization of functionals.

2.1 Calculus of Variations: a 101

Since taking variations will play an important part in Section 4, it is good to give a brief reminder of (or introduction to) the calculus of variations. The long and short of it is that, when *differentiating*, we are interested in how a quantity changes compared to an independent variable it depends on. When *taking a variation*, we look at how a quantity depending on a function – a so-called *functional* – changes when that "independent" function is slightly altered by means of a perturbation function – for an illustration, see Figure 2.1.

Since nothing beats mathematical notation, let us utilize some. Suppose we have a functional $J : \mathcal{Y} \to \mathbb{R} : y \mapsto J[y]$, defined by

$$J[y] = \int_{x_1}^{x_2} F(x, y(x), y'(x))\, dx. \tag{2.1}$$

Here, $\mathcal{Y}$ is the space of all allowed functions y. Then, when taking the variation of this function, we obtain the *functional derivative*, which is the topic of the following definition.

Definition 2.1. The quantity $\delta J/\delta y$ is called the *functional derivative of* J, and is defined by

$$\boxed{\int_{x_1}^{x_2} \frac{\delta J}{\delta y}\eta\, dx := \frac{d}{d\epsilon}\Big|_{\epsilon=0} \int_{x_1}^{x_2} F(x, y(x) + \epsilon\eta(x), y'(x) + \epsilon\eta'(x))\, dx.} \tag{2.2}$$

Here, η is a (small) perturbation to y which vanishes at the end points, i.e. $\eta(x_1) = \eta(x_2) = 0$. See Figure 2.1 for a visualization of η and its effect on y.

Notation 2.2. When writing $|_0$ we imply $|_{\epsilon=0}$.

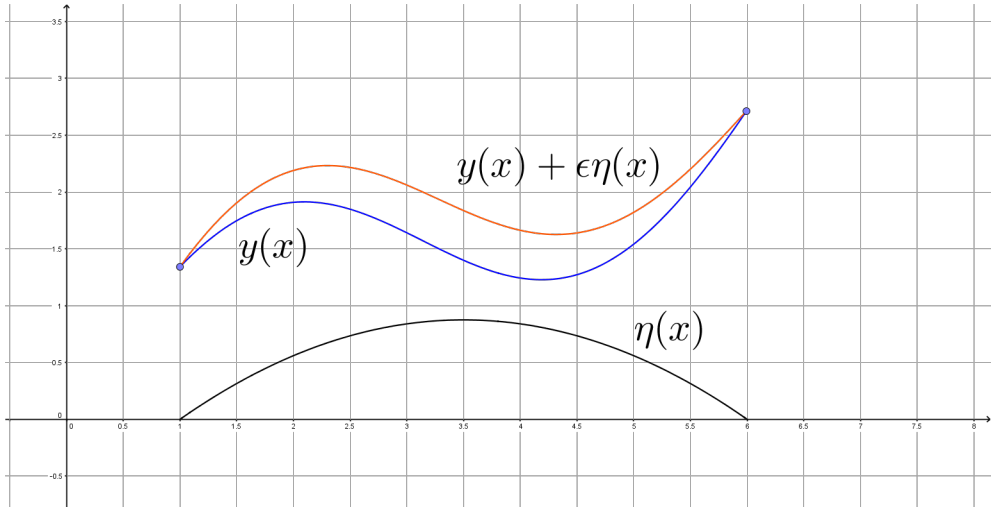


Figure 2.1: An illustration of the function y and its perturbation η. Notice how η is such that at the end points of the interval we wish to perturb, it vanishes.

Notation 2.3. From now on, the bounds on the integral signs will be omitted. However, this is done with the understanding that all integrals are still *definite* integrals.

Since Definition 2.1 might seem a little abstract, it is instructive to work a specific example of finding a functional derivative. Suppose we have a functional defined by

$$J = \int ay'y + b(y')^2x^2 - cy\,dx, \tag{2.3}$$

where a, b and c are constants and y depends on x, i.e. $y = y(x)$. Then, in staying with the definition, we do

$$\begin{aligned}
\int \frac{\delta J}{\delta y}\eta\,dx &= \frac{d}{d\epsilon}\Big|_0 \int a(y' + \epsilon\eta')(y + \epsilon\eta) + b(y' + \epsilon\eta')^2x^2 - c(y + \epsilon\eta)\,dx \\
&= \int \frac{\partial}{\partial\epsilon}\Big|_0 a\left[y'y + \epsilon(y\eta' + y'\eta) + \epsilon^2\eta\eta'\right] \\
&\quad + bx^2\left[(y')^2 + 2\epsilon y'\eta' + \epsilon^2(\eta')^2\right] - c[y + \epsilon\eta]\,dx \\
&= \int a(y\eta' + y'\eta) + 2bx^2y'\eta' - c\eta\,dx \\
&\overset{*}{=} \int 2bx^2\left((y'\eta)' - y''\eta\right) - c\eta\,dx \\
&\overset{*}{=} \int \eta\left(-2bx^2y'' - 4bxy' - c\right)\,dx \\
&\overset{**}{\Longrightarrow} \frac{\delta J}{\delta y} = -2bx^2y'' - 4bxy' - c.
\end{aligned} \tag{2.4}$$

In the above, the steps $*$ use integration by parts – to work the prime off of η – and the product rule.

The step $**$ is actually a little bit of a cheat. In general, it does *not* hold that whenever the integrals are equal, the integrands are equal as well. However, since our exclusive use of the variations is minimization (which is done by setting them to zero, as shown in the next section) the step $**$ effectively does hold – provided we also recognize that the entire expression should equal zero. As for why, this is because of the *fundamental lemma of the calculus of variations*, more information of which is provided in, for instance, [10].

The *very* astute mathematician will notice that the right-hand side of equation (2.4) is actually the Euler-Lagrange equation as applied to F with dependent variable y. This is a general result: the variation of a functional is given by the Euler-Lagrange equation, i.e.

$$\frac{\delta J}{\delta y} = \frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'}. \tag{2.5}$$

However, we will not be using the Euler-Lagrange equation. In Section 4 it will become clear that we will be required to take variations with respect to bras (in the Dirac formalism), a non-scalar object. Were we to use the Euler-Lagrange equation, it would require us to take *derivatives* with respect to said object, which would be awkward at best and incorrect at worst. As such, it is easiest (and mathematically safest) for us to stick to the definition of the variation as given in Definition 2.1.

One final point comes in the form of the chain rule.

Remark 2.4. We have that the chain rule holds for functional derivatives. That is, supposing we have two functionals

$$J_g = \int g(F(x, y, y'))\, dx, \text{ and } J_F = \int F(x, y, y')\, dx, \tag{2.6}$$

we have that

$$\frac{\delta J_g}{\delta y} = \frac{\partial g}{\partial F}\frac{\delta J_f}{\delta y}. \tag{2.7}$$

This can be seen by using the Euler-Lagrange equation and the chain rule for regular derivatives.

We shall require this later on. Namely, as functionals get more complicated (which they will), it will help greatly that we can do some simple differentiation prior to finding the functional derivative.

For further reference and background on functional derivatives, we refer the reader to [6].

2.2 Lagrangian multipliers

One notable feature of the classical brachistochrone problem covered in the introduction is that it was an *unconstrained* optimzation problem. That is, there were no restrictions on the function $y(x)$, i.e. $y(x)$ could be any function as long as it minimized the total travelling time.

However, this does not represent the totality of minimization problems that can be encountered. One example is ours: we shall see later on that we need to put restrictions on our system so that it represents a physical system (the Schrödinger equation comes to mind). A problem which has such a kind of restriction put on it is called a *constrained* optimization problem.

In general, unconstrained problems are much easier to work with than constrained problems (consider once more the classical brachistochrone: all we had to do was apply the Euler-Lagrange equation). It is therefore beneficial to somehow be able to rewrite any constrained problem into an unconstrained problem. This is the main use for the method of *Lagrange multipliers*.

The method goes as follows. Given are a quantity to minimize, $L(z), z \in \mathbb{R}^n$ and k constraints, formulated as

$$g_i(z) = 0 \quad \text{for } i = 1, 2, ..., k. \tag{2.8}$$

Then, the method of Lagrangian multipliers says that a minimizing solution to the *unconstrained* problem defined by

$$\overline{L}(z, \lambda) = L(z) + \sum_{i=1}^{k} \lambda_i g_i(z) \tag{2.9}$$

is also a solution to the constrained problem. In equation (2.9), $\lambda := (\lambda_1, \lambda_2, ..., \lambda_k)$ is called the *Lagrangian multiplier*. A full proof of this method can be found in various sources and textbooks, such as [10] that this thesis officially references.

The method as outlined above, though, treats the problem in a rather static manner. In particular, λ is assumed to be a constant vector and the minimizing solution is merely a single point. For our purposes, we would like to have a method based on the Lagrangian multiplier which handles functions as minimizing solutions, rather than single points. This will also necessitate ditching the assumption that λ is constant.

Fortunately, such a method exists. Suppose we have a functional we wish to optimize,

$$J[x, u] = \int_0^T L(x(t), u(t))\, dt, \tag{2.10}$$

where $u : [0, T] \to \mathbb{R}$ is some function in some function space $\mathcal{U}$ that we can *choose* so as to minimize J. Furthermore, x varies according to

$$\dot{x} = f(x, u). \tag{2.11}$$

Thus, we have a *constrained* minimization problem: J needs to be minimized whilst x *has to obey* the relation (2.11) at all time. But, since we prefer unconstrained problems, our Lagrangian multipliers senses start tickling. In that spirit, let us define the following quantity:

$$K(x, \dot{x}, p, u) := L(x, u) + p^T \cdot (f(x, u) - \dot{x}), \tag{2.12}$$

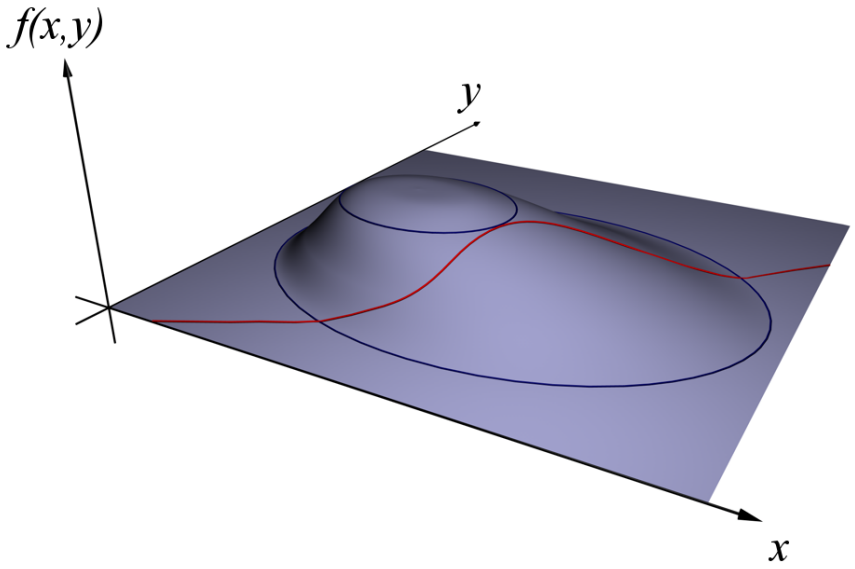


Figure 2.2: A graphical representation of Lagrange multipliers. The red line is the constraint, whilst the surface indicates the value of to be optimized function. As can be seen, the constraint is parallel to some level set of f where f is optimized. Source: khanacademy.com.

where $p : [0, T] \to \mathbb{R}^n$. Notice the similarity with Lagrangian multipliers: $f(x, u) - \dot{x} = 0$ is a constraint whilst p acts as Lagrange multiplier. Thus, the unconstrained problem associated with (2.10) is given as minimizing simply

$$J'[x, u] = \int_0^T K(x, \dot{x}, p, u)\, dt. \tag{2.13}$$

And we know how to minimize unconstrained problems: simply set the variations equal to zero. Setting all the variations equal to zero implies that whichever function we vary a little bit, the functional does not increase or decrease. [1] Therefore, in order to minimize our functional K, we have to look for those functions x, p and u which satisfy

$$\boxed{\frac{\delta K}{\delta x} = \frac{\delta K}{\delta p} = \frac{\delta K}{\delta u} = 0.} \tag{2.14}$$

We have now reduced finding the solution to the unconstrained problem of minimizing K to (simply) calculating some variations and determining for which functions they vanish. This is the strategy we will apply in Section 4 in order to find the minimizing solution of the quantum action, and so solve the quantum brachistochrone problem.

[1] It is like when finding the extrema of a function one looks at the (partial) derivatives and finds a point where they all vanish. If at some point they did not vanish, you could follow the derivative to go to a point which has lower value, implying that the point you are looking at is not the lowest.

There is one final note to give, though. Just as with derivatives it is imperative to check that the stationary point you have found is one you are looking for, so too it is with variations. Although there exist various tests to determine whether a minimum or maximum is found, we will not utilize these in this thesis, and assume that the solution we will find is the minimum solution.

For more information regarding Lagrangian multipliers – for instance a proof of their working – we refer the reader to [10].

3 Derivation of the action

This section concerns itself with finding the functional which describes the amount of time required to transition from one state in quantum space to another. This way, we can apply the calculus of variations as we learned it in Section 2.1 to optimize this functional by finding suitable state and Hamiltonian.

The functional we shall justify in this section is given below already. This is done so that we may associate each part of its structure to a particular subsection which will discuss it.

$$\begin{aligned} S[\psi, H, \phi, \lambda] = \int & \frac{\sqrt{\langle \partial_t \psi | (1-P) | \partial_t \psi \rangle}}{\Delta E} \\ & + \left(-i \langle \phi | \partial_t \psi \rangle + \langle \phi | H | \psi \rangle - i \langle \psi | \partial_t \phi \rangle + \langle \psi | H | \phi \rangle\right) \\ & + \lambda_1 (\mathrm{Tr}(\tilde{H}^2)/2 - \omega^2) + \sum_{j=2}^{m} \lambda_j \cdot f_j(H)\, dt, \end{aligned} \tag{3.1}$$

where ϕ and λ are Lagrange multipliers, and ω is a constant which can be interpreted as the energy uncertainty associated with the transition.[2]

Remark 3.1. In the original paper by Carlini *et al* [5], they instead use the nomenclature *action* to describe the functional (3.1) (whence also the symbol S). Why this is done specifically, I could not find out. However, in keeping with their naming, it has been adapted into this thesis.

Before we continue, there is a point that needs to be clarified: the bounds on the integral have been omitted for simplicity. As in Notation 2.3, though, we still consider the integrals to be definite by implicitly defining the bounds to be the initial and final states.

On to a brief summary of this section. In the first subsection, we will discuss what the first line of (3.1) represents: the quantity ds/v as phrased in quantum mechanical terms. In these terms, ds is the *quantum line element* on the space in which quantum states live, and v represents the "speed" at which they transition to other states. We shall derive the form of ds as grassroots as possible, by means of geometry on the sphere, and then work our way to the infinitesimal.

The second and third lines feature the constraints that our system has to obey – to name: the Schrödinger equation, the "finite energy" condition and miscellaneous constraints. These, as well as their incorporation into the integral by means of Lagrangian multipliers, will all be touched upon individually in the second part of this section.

3.1 The Quantum Line Element

In this section we will derive the main ingredient used in the quantum action (3.1): the infinitesimal time element dt associated with a certain transition. This way, in

[2]The energy uncertainty is *indeed* a constant, as we show further ahead in Lemma 4.8.

the spirit of the classical brachistochrone (equation (1.2)), we can then integrate to find the total time.

Since we are working in a finite dimensional Hilbert space, we consider our quantum states to be elements of $\mathbb{C}^n$, with additional structure provided by proportionality in $\mathbb{C}$. This yields the *complex projective space*, our first stop in this section.

We will derive the time element mainly by deriving the *quantum line element* ds – that is, the "distance" between two quantum states. In the first subsection, we will show that the aforementioned complex projective space can be thought of as a sphere with an equivalence relation. On this sphere, then, geometry will be conducted in order to find the distance between different quantum states – the *Fubini-Study distance*. Armed with a general formula, we then derive the infinitesimal form to obtain ds.

Definitions

Since we shall quite intensively use some geometrical concepts, it serves us well to give some definitions from the start.

Definition 3.2. The *norm* of a vector $X \in \mathbb{C}^n$ (or $X \in \mathbb{R}^n$) is given by

$$|X|^2 := X \cdot \overline{X} = \sum_{I=0}^{n} X^I \cdot \overline{X}^I, \tag{3.2}$$

where n is the dimension of the space and the overbar indicates complex conjugate (for real vectors, this is simply transposition).

Definition 3.3. The unit sphere embedded in $\mathbb{R}^{n+1}$ is $\mathbb{S}^n$, and is defined by

$$\mathbb{S}^n := \{X \in \mathbb{R}^{n+1} : |X|^2 = 1\}. \tag{3.3}$$

The sphere itself is n-dimensional, whence the superscript.[3]

Definition 3.4. The unit sphere embedded in $\mathbb{C}^{n+1}$ is $\mathbb{S}^{2n+1}$, defined as in Definition 3.3. The superscript still indicates the dimension of the sphere: this dimension is apparent from the observation that $\mathbb{C} \cong \mathbb{R}^2 \implies \mathbb{C}^n \cong \mathbb{R}^{2n}$.

3.1.1 Complex Projective Space

We start off by giving the definition of complex projective space.

Definition 3.5. The *complex projective space* $\mathbb{C}\mathbf{P}^n$ is the object

$$\mathbb{C}\mathbf{P}^n = \mathbb{C}^{n+1}/\sim, \tag{3.4}$$

where

$$x \sim y \quad \Longleftrightarrow \quad y = \lambda x, \text{ with } \lambda \in \mathbb{C}. \tag{3.5}$$

[3]For instance, the sphere embedded in $\mathbb{R}^3$ is a two-dimensional surface.

In layman's terms, this means that the equivalence class of some $z \in \mathbb{C}\mathbf{P}^n$ consists of all those points which are proportional to it, where proportionality is in $\mathbb{C}$. There are two things of particular importance we need to note.

1. The equivalence class of any point contains a point which has unity norm. This follows trivially from the observation that for any $z \in \mathbb{C}^{n+1} \setminus \{0\}$, there is $z' = z/|z| \in [z]$ which has $|z'| = 1$. Therefore, in complex projective space, $\mathbb{C}^{n+1}$ can be thought of as having been "brought back" to simply those elements with unity norm.
2. Membership of an equivalence class is invariant under phase change. That is, for any $z \in \mathbb{C}^{n+1}$, $z'' = e^{i\theta} z \in [z]$ for arbitrary θ. As such, all the points which lie on the same "great circle"[4] belong to the same equivalence class.

Hopefully, the above two observations convince the reader of the truth of the following theorem.

Theorem 3.6. *We have that*

$$\mathbb{C}\boldsymbol{P}^n = \mathbb{S}^{2n+1}/\mathbb{S}^1 := \{[x]_\sim : |x| = 1 \text{ for } x \in \mathbb{C}^{n+1}\}, \tag{3.6}$$

where $[x]_\sim = \{y \in \mathbb{C}^{n+1} \mid \exists\theta : y = e^{i\theta}x\}$.

Crucially, this allows us to think of the complex projective space as a sphere with equivalence between those elements differing by a phase.

This observation also justifies us investigating the complex projective space: it fits precisely with the quantum physicist's needs for a space, as i) it provides unity norm for all its elements, and ii) elements are equivalent under change in phase.

3.1.2 Geometry On Spheres

In the previous subsection we learned that, through $\mathbb{C}\mathbf{P}^n$, quantum states live on a spherical surface. This allows our quest for the quantum line element to be limited to geometry on spheres. In this section, we will seek to find an expression for the distance on spheres using geodesics.

Starting off our discussion is the notion of "distance" on a given surface (in our case a sphere).

Definition 3.7. The distance between two points is defined to be smallest path-length connecting both points. That is,

$$\text{distance} = \min_{\text{all paths } X} \text{length}(X) \tag{3.7}$$

heuristically. That path which minimizes the length (and hence yields the distance) is called a *geodesic*.

Thus, our problem is reduced to finding the geodesic on a sphere: once we know the geodesic, we simply find the length of that geodesic between two given points to find the distance.

[4] Great circle is in quotations as the analogy breaks down for higher dimensional spaces.

Since from Definition 3.7 we learn that finding the geodesic is essentially a minimization problem, we employ the calculus of variations. In this spirit, we propose the following functional which gives the length of a certain path X.

Lemma 3.8. *The functional to be minimized in order to find the geodesic is given by*

$$L[X] = \int F(\tau, X(\tau), X'(\tau))\, d\tau = \int \frac{1}{2}|X'|^2 + \lambda(|X|^2 - 1)\, d\tau \tag{3.8}$$

where $X(\tau)$ is a path on the sphere, the derivative is with respect to τ and λ is a Lagrange multiplier.

Proof. Consider the second term first. This is simply the constraint that $X \cdot X = 1$ for all τ, as should this not be fulfilled, X is no longer part of the sphere. Multiplied with this constraint is λ, in the spirit of Lagrange multipliers.

The first term dictates the thing we wish to minimize, namely $|X'|^2$. We wish to minimize this instead of the linear term, as this simplifies calculations down the line (Euler-Lagrange). □

We make one further assumption: that X is *parametrized by arclength*, or equivalently that $|X'|^2 = 1$. Being parametrized by arclength essentially means that the "time" defining the path (in our case τ) reflects the length of the path. This is a standard assumption/condition geometers put on their functions as it makes life easier, as it does for us.

We may then apply the Euler-Lagrange equation to (3.8). This yields:

$$X'' = -2\lambda X, \tag{3.9}$$

a second order differential equation which has as its solution

$$X^I(\tau) = k^I \cos(\tau) + \ell^I \sin(\tau), \quad \text{with } |k|^2 = |\ell|^2 = 1 \text{ and } k \cdot l = 0. \tag{3.10}$$

In equation (3.10), the vectors k and ℓ are constant and represent the initial position and direction of travel, respectively (through evaluating $X(0)$ and $X'(0)$). Furthermore, we used for the solution that $\lambda = 1/2$, a fact which follows from using that $|X| = |X'| = 1$.

Since X is parametrized by arclength, we have that the distance between points $X(\tau_1)$ and $X(\tau_2)$ is given by

$$\boxed{d = |\tau_1 - \tau_2|} \tag{3.11}$$

(this is a key element of what being parametrized by arclength entails; details can be found in textbooks on geometry). Notice that this is the first instance for which we have concretized the notion of distance on the sphere: we now know what we are talking about, as it were. We now present the main proposition of this section, relating distances to geodesics.

Lemma 3.9. *Let X be a geodesic on a sphere parametrized by arclength in τ. Then,*

$$\cos(d) = X(\tau_1) \cdot X(\tau_2), \tag{3.12}$$

for some τ_1 and τ_2, and d is as in equation (3.11).

Proof. We simply work out the multiplication.

$$\begin{aligned} X(\tau_1)\cdot X(\tau_2) &= (k\cos(\tau_1)+\ell\sin(\tau_1))(k\cos(\tau_2)+\ell\sin(\tau_2)) \\ &= k^2\cos(\tau_1)\cos(\tau_2)+\ell^2\sin(\tau_1)\sin(\tau_2)+k\cdot l(\cos(\tau_1)\sin(\tau_2)+\cos(\tau_2)\sin(\tau_1)) \\ &= \cos(\tau_1)\cos(\tau_2)+\sin(\tau_1)\sin(\tau_2) \\ &\overset{*}{=} \cos(|\tau_1-\tau_2|) \\ &= \cos(d), \end{aligned} \tag{3.13}$$

where for $*$ we used the trigonometric identity. □

Lemma 3.9, in effect, gives us a formula for finding the distance between points on the sphere. This simple formula will prove more than important as we advance to the next section

3.1.3 Adding complexity

The treatment of the previous subsection (and the results derived there) have concerned spheres embedded in real spaces. However, it will come as no surprise that there exist analogs of these results for spheres embedded in complex spheres. In particular, since our minimizing Lagrangian still holds, the solution also still holds – at least to some degree.

Proposition 3.10. *The geodesic on a sphere $\mathbb{S}^{2n+1}$ embedded in complex space is given by*

$$Z^\alpha(\tau) = m^\alpha\cos(\tau)+n^\alpha\sin(\tau), \tag{3.14}$$

where for $m^\alpha, n^\alpha \in \mathbb{C}^{n+1}$ it holds that

$$|m|^2 = |n|^2 = 1, \quad m\cdot\overline{n}+n\cdot\overline{m}=0. \tag{3.15}$$

Assuming moreover that Z is parametrized by arclength as well, we retain equation (3.11) and Lemma 3.9 transforms into the following.

Proposition 3.11. *Let Z be a geodesic on $\mathbb{S}^{2n+1}\subseteq\mathbb{C}^{n+1}$ parametrized by arclength in τ. Then,*

$$\cos(d) = \frac{1}{2}(Z(\tau_1)\overline{Z(\tau_2)}+Z(\tau_2)\overline{Z(\tau_1)}), \tag{3.16}$$

where

$$d = |\tau_1-\tau_2| \tag{3.17}$$

as in equation (3.11).

Proof. The proof involves the same working out as in the proof of Proposition 3.9, and is thus omitted. □

We have just entered the final stretch to finding the line element. Consider now the family of geodesics defined by

$$n^\alpha = im^\alpha \implies Z^\alpha(\tau) = m^\alpha\exp(i\tau), \tag{3.18}$$

of which

$$A^\alpha = A_0^\alpha \exp(i\tau), \quad B^\alpha = B_0^\alpha \exp(i(\tau + \tau_0)), \tag{3.19}$$

with $A_0, B_0 \in \mathbb{C}^{n+1}$ constant, are evidently members (τ_0 is a free parameter and will be of importance shortly). Note that although on the *sphere* these are two geodesics ("great cirles" of sorts), in the *complex projective space* it holds that

$$A^\alpha \sim A_0^\alpha, \quad B^\alpha \sim B_0^\alpha \tag{3.20}$$

for all τ. This way, if we consider the geodesics as living on $\mathbb{C}\mathbf{P}^n$, we are in effect looking at the distance between two points. Thus, we can use the formula as given in Proposition 3.11. Filling this in, we obtain

$$\begin{aligned}
\cos(d) = \frac{1}{2}(A \cdot \overline{B} + B \cdot \overline{A}) &= \frac{1}{2}(e^{i\tau} A_0 \cdot e^{-i(\tau+\tau_0)}\overline{B_0} + e^{i(\tau+\tau_0)} B_0 \cdot e^{-i\tau}\overline{A_0}) \\
&= \frac{1}{2}(A_0 \cdot \overline{B_0} e^{i\tau_0} + B_0 \cdot \overline{A_0} e^{i\tau_0}) \\
&\overset{*}{=} \frac{r}{2}(e^{i(\phi-\tau_0)} + e^{i(\tau_0-\phi)}) \\
&= r \cdot \cos(\phi - \tau_0),
\end{aligned} \tag{3.21}$$

where we took r and ϕ as defined by

$$A_0 \cdot \overline{B_0} = r e^{i\phi}. \tag{3.22}$$

This final step is legitimate, since $A_0, B_0 \in \mathbb{C}^{n+1}$, so that their inner product should give an element in $\mathbb{C}$, of which we have chosen the polar representation.

The obtained expression (3.21) still contains an unused τ_0, though, which we will utilize as follows. Since the distance between two points is the shortest possible path length between them, we can use τ_0 to minimize the length between A_0 and B_0.[5]

Since $\cos(d) \approx 1 - d^2/2$, a smallest possible value of d is accomplished by the highest possible value of $\cos(d)$. Thus, we are looking for the highest possible of $r\cos(\phi - \tau_0)$ – evidently r. The value of r can be obtained by choosing $\tau_0 = \phi$, so that $r\cos(\phi - \tau_0) = r\cos(\phi - \phi) = r$. Now dub that d which accomplishes $\cos(d) = r_0$ the *distance* d_0, i.e. we have that $\cos(d_0) = r$, where d_0 is the distance between two points on $\mathbb{C}\mathbf{P}^n$.

This measurement of distance on the complex projective space is known as the *Fubini-Study metric*, which is the most natural definition of distance we have on the complex projective space and, by extension, in the quantum world. As such, the quantum line element is in fact the infinitesimal form of the Fubini-Study metric. This is the topic of the following theorem, the apotheosis of our derivation.

Theorem 3.12. (Fubini-Study line element) *Let d_0 as has been defined earlier. Then, we have that infinitesimal version of the Fubini-Study metric is given by*

$$\boxed{ds^2 = dA \cdot d\overline{A} - dA \cdot \overline{A} A \cdot d\overline{A},} \tag{3.23}$$

when expanded up to second order in both d_0 and dA.

[5] Actually, with this step we are looking at the smallest length possible between all the elements in the equivalence classes A_0 and B_0, and the τ_0 that accomplishes that smallest length.

Proof. First of all, notice that we have

$$\cos^2(d_0) = r^2 = re^{i\phi} \cdot re^{-i\phi} = A_0 \cdot \overline{B_0} B_0 \cdot \overline{A_0} = A \cdot \overline{B} B \cdot \overline{A}. \tag{3.24}$$

We start by expanding the left-hand side of (3.24). We see that

$$\begin{aligned} \cos^2(d_0) &= \cos(0) + \frac{d[\cos^2(d_0)]}{dd_0}\Big|_0 ds + \frac{1}{2}\frac{d^2[\cos^2(d_0)]}{dd_0^2}\Big|_0 ds^2 + \mathcal{O}(ds^3) \\ &= 1 - 2\cos(0)\sin(0)ds - \frac{2\cos(2\cdot 0)}{2}ds^2 \\ &= 1 - ds^2, \end{aligned} \tag{3.25}$$

where we used ds as the infinitesimal version of d_0.

Expansion of the right-hand side of (3.24) is slightly trickier. First, we must slightly adjust the right-hand side slightly to read

$$\frac{A \cdot \overline{B}\, B \cdot \overline{A}}{A \cdot \overline{A}\, B \cdot \overline{B}} \tag{3.26}$$

instead. This is called the *projective cross ratio* κ. The format is quasi-justified by taking into account $A \cdot \overline{A} = B \cdot \overline{B} = 1$. The denominator is critical for the derivation, though, and so cannot be omitted.

Then, since we want to find the infinitesimal version of the distance, we effectively wish to find the distance between A and $A + \epsilon dA$, where $\epsilon = 1$.[6] We set $B := A + \epsilon dA$, so that the quantity to be expanded is thus

$$\frac{A \cdot (\overline{A} + \epsilon d\overline{A})\,(A + \epsilon dA)\overline{A}}{A \cdot \overline{A}\,(A + \epsilon dA) \cdot (\overline{A} + \epsilon d\overline{A})} =: \kappa(\epsilon) \tag{3.27}$$

with respect to epsilon. This results in

$$\begin{aligned} \kappa(\epsilon) &= \kappa(0) + \frac{d\kappa}{d\epsilon}\Big|_0 \epsilon + \frac{1}{2}\frac{d^2\kappa}{d\epsilon^2}\Big|_0 \epsilon^2 + \mathcal{O}(\epsilon^3) \\ &= 1 + 0 + \frac{1}{2}\frac{2(A \cdot d\overline{A}\, dA \cdot \overline{A} - A \cdot \overline{A}\, dA \cdot d\overline{A})}{(A \cdot \overline{A})^2}\epsilon^2 \\ &= 1 + \epsilon^2 \frac{A \cdot d\overline{A}\, dA \cdot \overline{A} - A \cdot \overline{A}\, dA \cdot d\overline{A}}{(A \cdot \overline{A})^2}. \end{aligned} \tag{3.28}$$

We make two final adjustments to equation (3.28), being i) we use $A \cdot \overline{A} = 1$, and ii) we set $\epsilon \equiv 1$, so that we are finding the distance between A and $A + dA$. This yields us

$$\kappa(1) = 1 + A \cdot d\overline{A}\, dA \cdot \overline{A} - dA \cdot d\overline{A}. \tag{3.29}$$

Finally, then, we put together equations (3.25) and (3.29) to find

$$1 - ds^2 = 1 + A \cdot d\overline{A}\, dA \cdot \overline{A} - dA \cdot d\overline{A}, \tag{3.30}$$

i.e.

$$ds^2 = dA \cdot d\overline{A} - A \cdot d\overline{A}\, dA \cdot \overline{A}, \tag{3.31}$$

proving the theorem.[7] □

[6] We find the infinitesimal distance this way so we can expand with respect to a scalar, instead of a vector.

[7] Do not forget to breathe.

3.1.4 Wrapping up

Theorem 3.12 thus yields us the quantum line element. However, we might prefer the following version of it, to comply with the standard – Dirac – notation of quantum mechanics.

Corollary 3.13. *The Fubini-Study line element as derived in Theorem 3.12 can be written in quantum mechanical notation as*

$$\boxed{ds^2 = \langle d\psi|(1-P)|d\psi\rangle\,,} \tag{3.32}$$

where $P = |\psi\rangle\langle\psi|$ *and* 1 *is the unit operator.*

Proof. Rewrite equation (3.23) using $A := \langle\psi|$, $\overline{A} = |\psi\rangle$, $dA = \langle d\psi|$ and $d\overline{A} = |d\psi\rangle$. □

We are *almost* there. The only ingredient we still require is the equation

$$\frac{ds}{dt} = \Delta E, \tag{3.33}$$

where $\Delta E := \langle H^2\rangle - \langle H\rangle^2$. This relation is shown in [1].[8] That is, in order to get an infinitesimal time element dt (over which we need to integrate in order to get the total time), we must have that

$$dt = \frac{ds}{\Delta E} = \frac{\sqrt{\langle d\psi|(1-P)|d\psi\rangle}}{\Delta E} = \frac{\sqrt{\langle \partial_t\psi|(1-P)|\partial_t\psi\rangle}}{\Delta E}\,dt. \tag{3.34}$$

Note that for the final equality sign, we "removed" a dt from the denominator, transforming $d\psi$ to $\partial_t\psi$. The equation (3.34) thus represents the infinitesimal time element, which we need to integrate in order to find the total time required for quantum states to change.

3.2 Constraints

This subsection is concerned with finding the constraints to which we apply the theory of Lagrange multipliers treated just now. The constraints we impose on the quantum state are that i) the Schrödinger equation is satisfied at all times, and ii) the energy uncertainty ΔE cannot be unbounded. Moreover, we allow for finitely more constraints to be imposed by means of a general formulation.

3.2.1 The Schrödinger equation

Since we are using Dirac notation for our quantum mechanics, we will use the Schrödinger equation in that format:

$$i\,|\partial_t\psi\rangle = H\,|\psi\rangle \quad\Longleftrightarrow\quad i\,|\partial_t\psi\rangle - H\,|\psi\rangle = 0. \tag{3.35}$$

[8] Actually, the article shows that $ds/dt = 2\Delta E$, but we assume that we can rescale in order to cancel the two.

Then, in the spirit of Lagrange multipliers, we left-multiply the right-hand side of the equivalence relation with the Lagrange multiplier $\langle\phi|$ to obtain

$$i\,\langle\phi|\partial_t\psi\rangle - \langle\phi|H|\psi\rangle = 0. \tag{3.36}$$

This is the term we will add to the action as contribution of the Schrödinger equation. Except, it is not the full picture. In order to fully capture the contribution, we must also consider the Hermitian conjugate of the Schrödinger equation, given by

$$-i\,\langle\partial_t\psi| = H\,\langle\psi| \quad\Longleftrightarrow\quad -i\,\langle\psi| - H\,\langle\psi| = 0. \tag{3.37}$$

In a similar spirit, now multiply with the suitable Lagrange multiplier to get

$$-i\,\langle\partial_t\psi|\phi\rangle - \langle\psi|H|\phi\rangle = 0. \tag{3.38}$$

Though usable, equation (3.38) is not the form we would prefer to use, as will become clear once we take variations in the next section. Luckily, we can rewrite using the following lemma.

Lemma 3.14. *The Hermitian conjugate of the time derivative operator is the negative of the time derivative operator, i.e.*

$$\partial_t^\dagger = -\partial_t. \tag{3.39}$$

Proof. Recall that $\langle\psi|\psi\rangle = 1$. Then,

$$\begin{aligned} 0 = \partial_t[\langle\psi|\phi\rangle] &= \partial_t \int \psi^*\phi\,dx \\ &= \int \partial_t\psi^*\phi + \psi^*\partial_t\phi\,dx \\ &= \langle\partial_t\psi|\phi\rangle + \langle\psi|\partial_t\phi\rangle \\ &\Longrightarrow \langle\psi|\partial_t\phi\rangle = \langle -\partial_t\psi|\phi\rangle\,. \end{aligned} \tag{3.40}$$

Thus, by definition of the Hermitian conjugate, $\partial_t^\dagger = -\partial_t$. □

In particular, we use Lemma 3.14 on the first term in equation (3.38) to obtain

$$-i\,\langle\psi|\partial_t\phi\rangle + \langle\psi|H|\phi\rangle = 0. \tag{3.41}$$

This term, in conjunction with equation (3.36), is what we will add to the action as representing the Schrödinger equation.

3.2.2 Boundedness of energy uncertainty

This condition is paramount in order to formulate a physically realistic system. For one, if we were to let the energy uncertainty grow arbitrarily, we could make the total action S arbitrarily small thanks to the appearance of ΔE in the denominator of the Fubini-Study metric.

As such an energy uncertainty could be obtained by suitable choice of Hamiltonian, it thus makes sense to impose a condition on H instead of ΔE. However,

the condition we elect to impose may seem to have simply fallen from sky. Namely, we impose

$$\mathrm{Tr}(\tilde{H}^2) = 2\omega^2, \tag{3.42}$$

where $\omega \in \mathbb{R}$ and $\tilde{H} := H - \mathrm{Tr}(H)/n$. A little rephrasing is in order to clarify what this condition means physically.

$$2\omega^2 = \mathrm{Tr}(\tilde{H}^2) = \mathrm{Tr}\left((H - \mathrm{Tr}(H)/n)^2\right) \tag{3.43}$$

The right-hand side represents, in some sense, the energy uncertainty of the system. Note that $\mathrm{Tr}(H)$ is the sum of all the eigenvalues of the Hamiltonian,[9] so that $\mathrm{Tr}(H)/n$ is the mean eigenvalue. Subtracting this from H so yields a Hamiltonian operator which has its eigenvalues downshifted, so that its new mean is zero. This way, the new eigenvalues represent a deviation from the mean. Then squaring this reduced Hamiltonian also squares all its eigenvalues, which are then added to each other by taking the trace once more. Thus, the "spread" of energies associated with the Hamiltonian is quantified, which in turn represents the energy uncertainty assocaited with the Hamiltonian.

We take the imposed constraint into account by using Lagrangian multipliers. Hence, the term we will add to our action to represent bounded energy uncertainty is

$$\lambda_1(\mathrm{Tr}(\tilde{H}^2)/2 - \omega^2). \tag{3.44}$$

3.2.3 Miscellaneous constraints

Despite the previous two constraints being important ones, these are not necessarily the only ones put on the system. For instance, there may be specific limitations to the equipment used in a laboratory setting, as such constraining the Hamiltonian operator acting on the quantum state. Or, in a quantum computer there are very specific voltages to work with so that only a restricted class of Hamiltonians can be allowed. In order to account for this, in this section we touch upon how further constraints can be added.

In particular, we consider only constraints imposed on the Hamiltonian; constraints on the state would be silly, as we can only indirectly affect it precisely through the Hamiltonian. Consider $n-1$ constraints phrased as

$$f_j(H) = 0, \quad \text{with } j = 2, 3, ..., n \tag{3.45}$$

and where $f_j : \mathcal{H} \to \mathbb{R} : H \mapsto f_j(H)$. In the spirit of Lagrangian multipliers, then, multiply all these functions with a specific λ and add them together to form the total constraint:

$$\sum_{j=2}^{n} \lambda_j f_j(H). \tag{3.46}$$

This is the the contribution of miscellaneous to the total action integral.

[9] Since H is a linear operator, this holds.

3.3 Final form

Collecting equations (3.34), (3.36), (3.41), (3.44) and (3.46) from the above subsections, we thus present the final form of the action to be

$$\begin{aligned} S(\psi, H, \phi, \lambda) = \int & \frac{\sqrt{\langle \partial_t \psi | (1-P) | \partial_t \psi \rangle}}{\Delta E} \\ & + (-i \langle \phi | \partial_t \psi \rangle + \langle \phi | H | \psi \rangle - i \langle \psi | \partial_t \phi \rangle + \langle \psi | H | \phi \rangle) \\ & + \lambda_1 (\mathrm{Tr}(\tilde{H}^2)/2 - \omega^2) + \sum_{j=2}^{m} \lambda_j \cdot f_j(H)\, dt. \end{aligned} \tag{3.47}$$

With this action in hand, we are able to derive the "equations of motion" for optimized transition between states, by taking variations with respect to all the different variables in play. This will the topic of the next section, where in addition we shall narrow down our constraints and as such arrive at the solutions of the quantum brachistochrone problem.

4 Variations, Equations Of Motion and The Solution

This section forms the heart of the thesis. Namely, here we will derive the equations of motion associated with the quantum action discussed in the previous section. These equations of motion will then be solved in order to obtain the optimal solution pair $|\psi\rangle, H$ which minimizes the transition time.

In the first subsection, we will take variations of our action. This way, in accordance with subsection 2.2, we then obtain the equations of motion.

Following this, the second subsection is concerned with solving these in the case of no additional constraints. That is, no constraints beyond the one outlined in subsection 3.2.2. This represents the ideal system, though not a realistic one.

The case where we *do* impose additional constraints is discussed in the final subsection of this section. It will come as no surprise that this will leave the most open-ended conclusion of the various subsections, as we cannot solve the system any further than we will without being given the constraints.

4.1 Taking various variations

Before we commence with taking variations, it is good to once more give the formula for the quantum action. This way, there will be no need for referencing to it in another section altogether.

$$\begin{aligned} S(\psi, H, \phi, \lambda) = \int & \underbrace{\frac{\sqrt{\langle \partial_t\psi|(1-P)|\partial_t\psi\rangle}}{\Delta E}}_{\text{(i)}} \\ & + \underbrace{(-i\langle\phi|\partial_t\psi\rangle + \langle\phi|H|\psi\rangle}_{\text{(ii)}} - \underbrace{i\langle\psi|\partial_t\phi\rangle + \langle\psi|H|\phi\rangle)}_{\text{(iii)}} \\ & \underbrace{+ \lambda_1(\mathrm{Tr}(\tilde{H}^2)/2 - \omega^2) + \sum_{j=2}^{m} \lambda_j \cdot f_j(H)\, dt.}_{\text{(iv)}} \end{aligned} \tag{4.1}$$

Then, we can commence with taking variations.

The first two are rather easy: the variations with respect to $\langle\phi|$ and λ. These are the subject of the following lemmas.

Lemma 4.1. *The variation of S with respect to $\langle\phi|$ is*

$$\frac{\delta S}{\delta\langle\phi|} = -i|\partial_t\psi\rangle + H|\psi\rangle. \tag{4.2}$$

Proof. Since only the (ii) term contains $\langle\phi|$, we can disregard the other terms and

focus solely on this one. We have that

$$\begin{aligned}\int \langle\delta\phi| \frac{\delta S}{\delta \langle\phi|}\, dt &= \frac{d}{d\epsilon}\Big|_0 \int -i \langle\phi + \epsilon\delta\phi|\partial_t\psi\rangle + \langle\psi + \epsilon\delta\phi|H|\psi\rangle \; dt \\ &= \int \langle\delta\phi| \left(-i\,|\partial_t\psi\rangle + H\,|\psi\rangle\right) dt \\ \implies \frac{\delta S}{\delta \langle\phi|} &= -i\,|\partial_t\psi\rangle + H\,|\psi\rangle\,. \end{aligned} \tag{4.3}$$

Thus, the lemma is proven. □

Lemma 4.2. *The variation of S with respect to λ is*

$$\frac{\delta S}{\delta\lambda} = \left(\frac{\delta S}{\delta\lambda_1}, \frac{\delta S}{\delta\lambda_2}, \ldots, \frac{\delta S}{\delta\lambda_m}\right) = \left(\frac{\mathrm{Tr}(\tilde{H}^2)}{2} - \omega^2, f_2(H), \ldots, f_m(H)\right) \tag{4.4}$$

Proof. We shall show that the lemma holds for each element separately, i.e. that $\delta S/\delta\lambda_k = f_k(H)$ for all $k = 1, 2, \ldots, m$ (and $f_1(H) = \mathrm{Tr}(\tilde{H}^2) - 2\omega^2$). The lemma then immediately follows.

Let it be seen that

$$\begin{aligned}\int \frac{\delta S}{\delta\lambda_k}\eta\, dt &= \frac{d}{d\epsilon}\Big|_0 \int (\lambda_k + \epsilon\eta) f_k(H)\, dt \\ &= \int \eta f_k(H)\, dt \\ \implies \frac{\delta S}{\delta\lambda_k} &= f_k(H). \end{aligned} \tag{4.5}$$

Thus, since the above holds for arbitrary k, it holds for all k, and so the lemma is proven. □

Notice that the previous two lemmas imply that the constraints we imposed on our system have to hold for any optimal solution $|\psi\rangle$ and H – i.e. which satisfy $\delta S/\delta\langle\phi| = \delta S/\delta\lambda = 0$. Thus, from now on we can effectively assume that $|\psi\rangle$ and H fulfill

$$i\,|\partial_t\psi\rangle = H\,|\psi\rangle\,, \quad \mathrm{Tr}(H)/2 = \omega^2, \quad f_j(H) = 0 \tag{4.6}$$

for $j = 2, 3, ..., m$. Here we recognize the power of the Lagrange multipliers, now brought out of theory and into practice. By working the constraints into the functional by means of Lagrange multipliers, they are now part of the equations of motion for which we have to solve to obtain an optimal solution *instead* of being separate constraints we would have had to consider.

The variations with respect to $\langle\psi|$ and H are a little trickier, though, and involve more mathematical subtleties. We acknowledge that the derivations presented here find their inspiration in [8]. Before we handle these, first some additional notation.

Notation 4.3. Upon defining a *function* G, the associated *functional* is denoted

$$R_G := \int G\, dt. \tag{4.7}$$

Commencing with taking the variation w.r.t. $\langle\psi|$, we have the following proposition.

Proposition 4.4. *The variation of S with respect to $\langle\psi|$ is*

$$\frac{\delta S}{\delta\langle\psi|} = i\partial_t\left[\frac{H-\langle H\rangle}{2(\Delta E)^2}\right]|\psi\rangle - i\,|\partial_t\phi\rangle + H\,|\phi\rangle \tag{4.8}$$

Proof. Since the terms (ii) and (iv) in equation (4.1) do not contain a $\langle\psi|$, we can disregard these for taking the variation.

(i): Define

$$A := \langle\partial_t\psi|(1-P)|\partial_t\psi\rangle\,,\; B := (\Delta E)^2 \equiv \langle H^2\rangle - \langle H\rangle^2 \tag{4.9}$$

so that, in effect, we need to find $\delta R_{\sqrt{A/B}}/\delta\langle\psi|$. We first utilize the chain rule – Remark 2.4 – so as to simplify:

$$\begin{aligned}\frac{\delta R_{\sqrt{A/B}}}{\delta\langle\psi|} &= \frac{1}{2}\sqrt{\frac{B}{A}}\cdot\frac{\delta R_{A/B}}{\delta\langle\psi|}\\ &= \frac{1}{2}\sqrt{\frac{B}{A}}\left(\frac{1}{B}\frac{\delta R_A}{\delta\langle\psi|} - \frac{A}{B^2}\frac{\delta R_B}{\delta\langle\psi|}\right)\\ &= \frac{1}{2}\sqrt{\frac{1}{AB}}\frac{\delta R_A}{\delta\langle\psi|} - \frac{1}{2B}\sqrt{\frac{A}{B}}\frac{\delta R_B}{\langle\psi|},\end{aligned} \tag{4.10}$$

The next step is thus to find δA and δB, of which we will treat the former first. Taking to heart the definition of taking a variation, we calculate

$$\begin{aligned}\int\langle\delta\psi|\frac{\delta R_A}{\delta\langle\psi|}\,dt &= \frac{d}{d\epsilon}\Big|_0\int\left[\,\langle\partial_t\psi+\epsilon\partial_t\delta\psi|(1-|\psi\rangle\langle\psi+\epsilon\delta\psi|)|\partial_t\psi\rangle\,\right]dt\\ &= \int\frac{d}{d\epsilon}\Big|_0\left[\,\langle\partial_t\psi+\epsilon\partial_t\delta\psi|\partial_t\psi\rangle\,\right] + \frac{d}{d\epsilon}\Big|_0\left[\,\langle\partial_t\psi+\epsilon\partial_t\delta\psi|\psi\rangle\;\langle\psi+\epsilon\delta\psi|\partial_t\psi\rangle\,\right]dt\\ &= \int\langle\partial_t\delta\psi|\partial_t\psi\rangle - \langle\partial_t\psi+\epsilon\partial_t\delta\psi|\psi\rangle\;\langle\delta\psi|\partial_t\psi\rangle\,|_{\epsilon=0}\\ &\quad - \langle\partial_t\delta\psi|\psi\rangle\;\langle\psi+\epsilon\delta\psi|\partial_t\psi\rangle\,|_{\epsilon=0}\,dt\\ &= \int\langle\partial_t\delta\psi|\partial_t\psi\rangle - \langle\partial_t\psi|\psi\rangle\;\langle\delta\psi|\partial_t\psi\rangle - \langle\partial_t\delta\psi|\psi\rangle\;\langle\psi|\partial_t\psi\rangle\;dt\\ &= \int\langle\partial_t\delta\psi|(1-P)|\partial_t\psi\rangle - \langle\delta\psi|\partial_t\psi\rangle\;\langle\partial_t\psi|\psi\rangle\;dt\\ &= \int\langle\delta\psi|\left(-\partial_t\left\{(1-P)\,|\partial_t\psi\rangle\right\} - \langle\partial_t\psi|\psi\rangle\;|\partial_t\psi\rangle\,\right)dt\\ \implies \frac{\delta R_A}{\delta\langle\psi|} &= -\partial_t\left\{(1-P)\,|\partial_t\psi\rangle\right\} - \langle\partial_t\psi|\psi\rangle\;|\partial_t\psi\rangle\,.\end{aligned} \tag{4.11}$$

A similar procedure for B yields

$$\begin{aligned}\int \langle\delta\psi|\,\frac{\delta R_B}{\delta\langle\psi|}\,dt &= \frac{d}{d\epsilon}\Big|_0 \int \langle\psi+\epsilon\delta\psi|H^2|\psi\rangle - (\,\langle\psi+\epsilon\delta\psi|H|\psi\rangle)^2\;dt \\ &= \int \langle\delta\psi|H^2|\psi\rangle - 2\,\langle\psi+\epsilon\delta\psi|H|\psi\rangle\;\langle\delta\psi|H|\psi\rangle\,|_{\epsilon=0}\,dt \\ &= \int \langle\psi|\left(H^2\,|\psi\rangle - 2\,\langle H\rangle\,H\,|\psi\rangle\right)\,dt \\ \implies \frac{\delta R_B}{\delta\langle\psi|} &= H^2\,|\psi\rangle - 2\,\langle H\rangle\,H\,|\psi\rangle\,. \end{aligned} \tag{4.12}$$

Now that we have taken variations of A and B, we can set their values to be $A = B = (\Delta E)^2$.[10] Then, combining the expressions (4.10), (4.11) and (4.12), assuming that ΔE is constant[11] (and using the Schrödinger equation a bunch of times), we obtain

$$\begin{aligned}\frac{\delta R_{\sqrt{A/B}}}{\delta\langle\psi|} &= \frac{1}{2(\Delta E)^2}\left(\frac{\delta R_A}{\delta\langle\psi|} - \frac{\delta R_B}{\delta\langle\psi|}\right) \\ &= \frac{1}{2(\Delta E)^2}\left(-\partial_t\left\{(1-P)\,|\partial_t\psi\rangle\right\} - \langle\partial_t\psi|\psi\rangle\,|\partial_t\psi\rangle - H^2\,|\psi\rangle + 2\,\langle H\rangle\,H\,|\psi\rangle\right) \\ &= \frac{1}{2(\Delta E)^2}\left(i\partial_t\left\{(1-P)H\,|\psi\rangle\right\} - \langle H\rangle\,H\,|\psi\rangle - H^2\,|\psi\rangle + 2\,\langle H\rangle\,H\,|\psi\rangle\right) \\ &= \frac{1}{2(\Delta E)^2}\left(i\partial_t\left\{[H-\langle H\rangle]\,|\psi\rangle\right\} + \langle H\rangle\,H\,|\psi\rangle - H^2\,|\psi\rangle\right) \\ &= \frac{1}{2(\Delta E)^2}\left(i\partial_t\left\{H-\langle H\rangle\right\}|\psi\rangle + [H^2 - \langle H\rangle\,H]\,|\psi\rangle + \langle H\rangle\,H\,|\psi\rangle - H^2\,|\psi\rangle\right) \\ &= i\partial_t\left[\frac{H-\langle H\rangle}{2(\Delta E)^2}\right]|\psi\rangle\,, \end{aligned} \tag{4.13}$$

which concludes the variation of term (i).

(iii): We set $I := \langle\psi|\partial_t\phi\rangle + \langle\psi|H|\phi\rangle$. Then, we have that

$$\begin{aligned}\int \langle\delta\psi|\,\frac{\delta R_I}{\delta\langle\psi|}\,dt &= \frac{d}{d\epsilon}\Big|_0 \int -i\,\langle\psi+\epsilon\delta\psi|\partial_t\phi\rangle + \langle\psi+\epsilon\delta\psi|H|\phi\rangle\;dt \\ &= \int \langle\delta\psi|\,(-i\,|\partial_t\phi\rangle + H\,|\phi\rangle)\;dt \\ \implies \frac{\delta R_I}{\delta\langle\psi|} &= -i\,|\partial_t\phi\rangle + H\,|\phi\rangle\,. \end{aligned} \tag{4.14}$$

We then combine the equations (4.13) and (4.14) so that we obtain the full variation, being

$$\frac{\delta S}{\delta\langle\psi|} = i\partial_t\left[\frac{H-\langle H\rangle}{2(\Delta E)^2}\right]|\psi\rangle - i\partial_t\,|\psi\rangle + H\,|\psi\rangle\,. \tag{4.15}$$

[10] For A, this is justified upon assuming the Schrödinger equation to hold and working out the original expression.

[11] This is not a trivial assumption, as H is assumed to be time-dependent. However, as is shown in Lemma 4.8, $(\Delta E)^2$ is constant indeed. Since the lemma does not use $\delta S/\delta\langle\psi|$, this is thus a consistent assumption to make.

This is precisely what we are looking for, so the proposition is proven. □

And now, the variation with respect to H.

Proposition 4.5. *The variation of S as in equation (4.1) with respect to H is given by*

$$\frac{\delta S}{\delta H} = \frac{2\langle H\rangle P - \{H,P\}}{2(\Delta E)^2} + |\psi\rangle\langle\phi| + |\phi\rangle\langle\psi| + \lambda_1\tilde{H} + \sum_{j=2}^{m}\lambda_j\frac{\delta f_j}{\delta H}. \tag{4.16}$$

Here, $\{H,P\} := HP + PH$ denotes the anticommutator between H and P.

Proof. We follow once more the lead of [8], where we now look at the derivation of equation (65). Looking at equation (4.1), we see that all terms involve an H, so that we need to consider all terms in taking this variation. In the end, we add up all the variations to get the grand total.

(i) As with Proposition 4.4, we choose $A := \langle\partial_t\psi|(1-P)|\partial_t\psi\rangle$ and $B := (\Delta E)^2$. Then,

$$\frac{\delta R_{\sqrt{A/B}}}{\delta H} = \frac{-1}{2B}\sqrt{\frac{A}{B}}\frac{\delta R_B}{\delta H}. \tag{4.17}$$

Notice that $\delta R_A/\delta H = 0$ as A does not (explicitly) depend on H. Then, we determine

$$\begin{aligned}\int \delta H\frac{\delta R_B}{\delta H} &= \frac{d}{d\epsilon}\Big|_0\int \left\langle (H+\epsilon\delta H)^2\right\rangle - \langle H+\epsilon\delta H\rangle^2\, dt\\ &= \int \frac{d}{d\epsilon}\Big|_0\left\langle H^2 + \epsilon(H\delta H + \delta H H) + \epsilon^2(\delta H)^2\right\rangle - \frac{d}{d\epsilon}\Big|_0\langle H+\epsilon\delta H\rangle^2\, dt\\ &= \int \langle H\delta H + \delta H H\rangle - 2\langle H\rangle\langle\delta H\rangle\, dt\\ \implies \frac{\delta R_B}{\delta H} &= \frac{\langle H\delta H\rangle}{\delta H} + \frac{\langle\delta H H\rangle}{\delta H} - 2\langle H\rangle\frac{\langle\delta H\rangle}{\delta H},\end{aligned} \tag{4.18}$$

so that we now effectively have to look at three different quantities: $\langle\delta H\rangle$, $\langle H\delta H\rangle$ and $\langle\delta H H\rangle$. Let us consider the former of the triple first. Expanding in an arbitrary basis $\{\alpha_i\}_{i=1}^n$, we have that

$$\langle\delta H\rangle = \langle\psi|\delta H|\psi\rangle = \langle\psi|\alpha_j\rangle\langle\alpha_j|\delta H|\alpha_k\rangle\langle\alpha_k|\psi\rangle, \tag{4.19}$$

where summation over j and k is implied.[12] Then, dubbing $\delta H_{jk} := \langle\alpha_j|\delta H|\alpha_k\rangle$, we obtain

$$\langle\psi|\alpha_j\rangle\langle\alpha_j|\delta H|\alpha_k\rangle\langle\alpha_k|\psi\rangle = \delta H_{jk}\langle\alpha_k|\psi\rangle\langle\psi|\alpha_j\rangle = \delta H_{jk}P_{kj}, \tag{4.20}$$

where, recall, $P := |\psi\rangle\langle\psi|$. As such, we have that

$$\frac{\langle\delta H\rangle}{\delta H_{jk}} = P_{kj} \implies \frac{\langle\delta H\rangle}{\delta H} = P. \tag{4.21}$$

[12] Otherwise known as the Einstein summation convention.

One down, two to go. Consider now $\langle \delta H H \rangle$. We proceed similarly as before:

$$\langle \delta H H \rangle = \langle \psi | \delta H H | \psi \rangle = \langle \psi | \alpha_j \rangle \langle \alpha_j | \delta H | \alpha_k \rangle \langle \alpha_k | H \psi \rangle . \tag{4.22}$$

Then,

$$\langle \psi | \alpha_j \rangle \langle \alpha_j | \delta H | \alpha_k \rangle \langle \alpha_k | H \psi \rangle = \delta H_{jk} \langle \alpha_k | H \psi \rangle \langle \psi | \alpha_j \rangle = \delta H_{jk} (HP)_{kj}, \tag{4.23}$$

resulting in

$$\frac{\langle \delta H H \rangle}{\delta H_{jk}} = (HP)_{kj} \implies \frac{\langle \delta H H \rangle}{\delta H} = HP. \tag{4.24}$$

In a similar way, it can be determined that

$$\frac{\langle H \delta H \rangle}{\delta H} = PH. \tag{4.25}$$

Then, joining equations (4.21), (4.24) and (4.25) with (4.12) yields

$$\frac{\delta B}{\delta H} = PH + HP - 2 \langle H \rangle P, \tag{4.26}$$

so that (4.17) becomes (recall that $A = B = (\Delta E)^2$)

$$\frac{\delta \left[\sqrt{A/B} \right]}{\delta H} = \frac{2 \langle H \rangle P - (HP + PH)}{2(\Delta E)^2} = \frac{2 \langle H \rangle P - \{H, P\}}{2(\Delta E)^2}, \tag{4.27}$$

the contribution of (i) to the total variation.

(ii)&(iii): We first direct our attention to $\langle \phi | H | \psi \rangle$, leading us to determine $\langle \phi | \delta H | \psi \rangle$. As we did for (i), we expand in a basis $\{\alpha_i\}_{i=1}^n$, obtaining

$$\langle \phi | \delta H | \psi \rangle = \langle \phi | \alpha_j \rangle \langle \alpha_j | \delta H | \alpha_k \rangle \langle \alpha_k | \psi \rangle = \delta H_{jk} \langle \alpha_k | \psi \rangle \langle \phi | \alpha_j \rangle . \tag{4.28}$$

Therefore,

$$\frac{\langle \phi | \delta H | \psi \rangle}{\delta H_{jk}} = (|\psi\rangle\langle\phi|)_{kj} \implies \frac{\langle \phi | \delta H | \psi \rangle}{\delta H} = |\psi\rangle\langle\phi| . \tag{4.29}$$

Similarly,

$$\frac{\langle \psi | \delta H | \phi \rangle}{\delta H} = |\phi\rangle\langle\psi| . \tag{4.30}$$

(iv): For the cases where $j > 1$, this is moot; we cannot take variations of functions we do not know. Thus, the best we can do is

$$\frac{\delta}{\delta H} \sum_{j=2}^{m} \lambda_j f_j(H) = \sum_{j=2}^{m} \lambda_j \frac{\delta f_j}{\delta H}. \tag{4.31}$$

For $j = 1$, we can do something. First off, note that

$$\frac{\lambda_1 \, \delta \operatorname{Tr}(\tilde{H}^2)}{2} = \frac{\lambda_1 \operatorname{Tr}(\delta[\tilde{H}^2])}{2} = \frac{\lambda_1}{2} \Big(\operatorname{Tr}(\tilde{H} \delta \tilde{H}) + \operatorname{Tr}(\delta \tilde{H} \tilde{H}) \Big) . \tag{4.32}$$

Then, the definition of the trace (sum of diagonal elements) allows us to rewrite:

$$\frac{\lambda_1}{2}\Big(\mathrm{Tr}(\tilde{H}\delta\tilde{H}) + \mathrm{Tr}(\delta\tilde{H}\tilde{H})\Big) = \frac{\lambda_1}{2}\Big[\tilde{H}_{ji}\delta\tilde{H}_{ij} + \delta\tilde{H}_{ij}\tilde{H}_{ji}\Big] = \lambda_1\tilde{H}_{ji}\delta\tilde{H}_{ij}. \tag{4.33}$$

As a consequence,

$$\frac{\lambda_1}{2}\frac{\delta\,\mathrm{Tr}(\tilde{H}^2)}{\delta\tilde{H}_{ij}} = \lambda_1\tilde{H}_{ji} \implies \frac{\lambda_1}{2}\frac{\delta\,\mathrm{Tr}(\tilde{H}^2)}{\delta H} = \lambda_1\tilde{H}, \tag{4.34}$$

where in the last step switching from $\delta\tilde{H}$ to δH is justified as . Then, collecting everything, we get

$$\lambda_1\tilde{H} + \sum_{j=2}^{m}\lambda_j\frac{\delta f_j}{\delta H} \tag{4.35}$$

as the contribution of (iv) to the total variation.

Thus, upon combining equations (4.27), (4.29), (4.30) and (4.35), we find that the total variation of S with respect to H becomes

$$\frac{\delta S}{\delta H} = \frac{2\langle H\rangle P - \{H,P\}}{2(\Delta E)^2} + |\psi\rangle\langle\phi| + |\phi\rangle\langle\psi| + \lambda_1\tilde{H} + \sum_{j=2}^{m}\lambda_j\frac{\delta f_j}{\delta H}, \tag{4.36}$$

which proves the proposition. □

Now that we have taken all the variations, it is time to utilize the fact that we are working with an unconstrained problem (in the sense that we have worked all the constraints into the extremand by means of Lagrange multipliers). Since our problem is unconstrained, the minimizing solution can be found by simply setting all the variations equal to zero, i.e. our solution pair $|\psi\rangle, H$ has to ensure that

$$\frac{\delta S}{\delta\langle\phi|} = \frac{\delta S}{\delta\lambda} = \frac{\delta S}{\delta\langle\psi|} = \frac{\delta S}{\delta H} = 0 \tag{4.37}$$

is satisfied. The former two of these we already assumed to hold; these are the constraints of our system being fulfilled. Thus, what is left for us to solve are the latter two equations. That is, we need to solve

$$\boxed{i\partial_t\left[\frac{H-\langle H\rangle}{2(\Delta E)^2}\right]|\psi\rangle - i\,|\partial_t\phi\rangle + H\,|\phi\rangle = 0,} \tag{4.38}$$

and

$$\boxed{\frac{2\langle H\rangle P - \{H,P\}}{2(\Delta E)^2} + |\psi\rangle\langle\phi| + |\phi\rangle\langle\psi| + \lambda_1\tilde{H} + \sum_{j=2}^{m}\lambda_j\frac{\delta f_j}{\delta H} = 0.} \tag{4.39}$$

Thus, recapping: as a result of taking variations, we have found the various equations of motion of the quantum brachistochrone system. Our next task, naturally, shall be to solve these, which is the topic of the following subsections.

4.2 Case with no constraints

At this point, it serves us to consider the case where we impose solely the trivial constraint separately from when we have m different constraints. The reason is that, when no additional constraints but the trivial one is imposed, the problem is simpler (less constraints means less stress), still somewhat represents a physical system and can be solved in closed form. Moreover, its derivation is very instructive, also with an eye to when additional constraints *are* imposed.

Thus, from now on we treat $f_j \equiv 0$ for $j = 2, 3, ..., m$. Immediately, we notice a simplification of equation (4.16):

$$\frac{\delta S}{\delta H} = \frac{2\langle H\rangle P - \{H, P\}}{2(\Delta E)^2} + (\,|\psi\rangle\langle\phi| + |\phi\rangle\langle\psi|\,) + \lambda\tilde{H}, \tag{4.40}$$

where we wrote $\lambda \equiv \lambda_1$ for convenience. Thus, the equations we need to solve for $\tilde{H}$ and $|\psi\rangle$ become

$$\boxed{i\partial_t\left[\frac{H - \langle H\rangle}{2(\Delta E)^2}\right]|\psi\rangle - i\partial_t|\phi\rangle + H|\phi\rangle = 0} \tag{4.41}$$

and

$$\boxed{\frac{\{H, P\} - 2\langle H\rangle P}{2(\Delta E)^2} - \lambda\tilde{H} - (\,|\psi\rangle\langle\phi| + |\phi\rangle\langle\psi|\,) = 0.} \tag{4.42}$$

Our strategy in solving these shall be two-fold. First, we will reduce the equations to a simpler form by performing extensive algebra. Secondly, we will solve these simplified equations in order to find the solution for the quantum brachistochrone. These are the topics of the following two subsections, respectively.

4.2.1 Setting Up The Equations

In this section, we shall aim to simplify the found expressions for the two variations. Namely, in their current state, the equations are hard and some work can be done in order to simplify.

Our first aim shall be to find an expression for $|\phi\rangle$. Then, we shall see that we can plug this back in and so simplify both variations into more manageable formats.

Before we commence with that, though, we present a triplet of lemmas which we will need for our derivations.

Lemma 4.6. *We have that $\langle\psi|\phi\rangle$ is purely imaginary, i.e.*

$$\langle\psi|\phi\rangle = -\langle\phi|\psi\rangle\,. \tag{4.43}$$

Proof. Taking the trace of equation (4.42) yields

$$\frac{\mathrm{Tr}(HP) + \mathrm{Tr}(PH) - 2\langle H\rangle\,\mathrm{Tr}(P)}{2(\Delta E)^2} - \lambda\,\mathrm{Tr}(\tilde{H}) - \mathrm{Tr}\,(\,|\psi\rangle\langle\phi|\,) - \mathrm{Tr}\,(\,|\phi\rangle\langle\psi|\,) = 0. \tag{4.44}$$

But, using the facts that

$$\mathrm{Tr}(HP) = \mathrm{Tr}(HP^2) = \mathrm{Tr}(PHP) = \mathrm{Tr}(\,|\psi\rangle\langle\psi|H|\psi\rangle\langle\psi|\,) = \langle H\rangle\,\mathrm{Tr}(P), \tag{4.45}$$

ditto for $\mathrm{Tr}(PH)$ and

$$\mathrm{Tr}(\tilde{H}) = \mathrm{Tr}(H) - \mathrm{Tr}(\mathrm{Tr}(H)/n) = \mathrm{Tr}(H) - \mathrm{Tr}(H)/n \cdot n = 0, \tag{4.46}$$

we as such have that equation (4.44) simplifies to

$$\mathrm{Tr}\,(\,|\psi\rangle\langle\phi|\,) = -\,\mathrm{Tr}\,(\,|\phi\rangle\langle\psi|\,)\,. \tag{4.47}$$

Now,

$$\mathrm{Tr}\,(\,|\psi\rangle\langle\phi|\,) = \mathrm{Tr}\,(P\,|\psi\rangle\langle\phi|\,) = \mathrm{Tr}\,(\,|\psi\rangle\langle\phi|\,P) = \langle\,|\psi\rangle\langle\phi|\,\rangle = \langle\phi|\psi\rangle\,, \tag{4.48}$$

and ditto for $\mathrm{Tr}\,(\,|\phi\rangle\langle\psi|\,)$. Therefore,

$$\mathrm{Tr}\,(\,|\psi\rangle\langle\phi|\,) = -\,\mathrm{Tr}\,(\,|\phi\rangle\langle\psi|\,) \implies \langle\psi|\phi\rangle = -\,\langle\phi|\psi\rangle\,. \tag{4.49}$$

The last statement proves the lemma. □

Lemma 4.7. *We have that* $\langle\tilde{H}\rangle = 0$*, from which it immediately follows that*

$$\langle H\rangle = \mathrm{Tr}(H)/n. \tag{4.50}$$

Proof. We take the expectation w.r.t $|\psi\rangle$ of equation (4.42), yielding

$$\left\langle \frac{\{H,P\} - 2\,\langle H\rangle\,P}{2(\Delta E)^2} \right\rangle - \lambda\,\langle\tilde{H}\rangle - (\,\langle\,|\psi\rangle\langle\phi|\,\rangle + \langle\,|\phi\rangle\langle\psi|\,\rangle) = 0. \tag{4.51}$$

The first and third terms cancel (by definition of P, the expectation and Lemma 4.6), so that remains

$$\lambda\,\langle\tilde{H}\rangle = 0 \implies \langle\tilde{H}\rangle = 0. \tag{4.52}$$

From the definition of $\tilde{H}$ it then readily follows that

$$0 = \langle\tilde{H}\rangle = \langle H\rangle - \langle\mathrm{Tr}(H)/n\rangle = \langle H\rangle - \mathrm{Tr}(H)/n \implies \langle H\rangle = \mathrm{Tr}(H)/n, \tag{4.53}$$

proving the lemma. □

Lemma 4.8. *We have that*

$$(\Delta E)^2 = \langle\tilde{H}^2\rangle = \mathrm{Tr}(\tilde{H}^2)/2 = \omega^2; \tag{4.54}$$

notably, $(\Delta E)^2$ *is constant.*

Proof. The first equality holds, as

$$\langle\tilde{H}^2\rangle = \langle H^2 + \langle H\rangle^2 - 2\,\langle H\rangle\,H\rangle = \langle H^2\rangle - \langle H\rangle^2 = (\Delta E)^2. \tag{4.55}$$

In justifying the second equality, we use that

$$\langle\tilde{H}^2\rangle = \mathrm{Tr}(\tilde{H}^2P) = \mathrm{Tr}(\tilde{H}P\tilde{H}) \tag{4.56}$$

by means of

$$\mathrm{Tr}(\tilde{H}P\tilde{H}) \overset{*}{=} \mathrm{Tr}(\tilde{H}^2 - \tilde{H}P\tilde{H}) = \mathrm{Tr}(\tilde{H}^2) - \mathrm{Tr}(\tilde{H}P\tilde{H}) \tag{4.57}$$

where $*$ holds by Proposition 4.9.[13][14] Therefore,

$$2\,\mathrm{Tr}(\tilde{H}P\tilde{H}) = \mathrm{Tr}(\tilde{H}^2) \implies \langle\tilde{H}^2\rangle = \mathrm{Tr}(\tilde{H}P\tilde{H}) = \mathrm{Tr}(\tilde{H}^2)/2. \tag{4.58}$$

And, by constraint, the latter equality holds:

$$\mathrm{Tr}(\tilde{H}^2)/2 = \omega^2. \tag{4.59}$$

As such, the lemma is proven. □

Using the above lemmas, then, we can proceed with simplifying the equations. Since we have two equations to begin with, we also need to end up with two other equations to solve. This will be the topic of the two following propositions.

Proposition 4.9. *Equation* (4.42) *implies that instead of solving* it*, we can equivalently solve*

$$\tilde{H} = \tilde{H}P + P\tilde{H}. \tag{4.60}$$

Proof. Apply $|\psi\rangle$ to (4.42) to obtain

$$\begin{aligned} 0 &= \left[\frac{\{H,P\} - 2\langle H\rangle P}{2(\Delta E)^2} - \lambda\tilde{H} - (\,|\psi\rangle\langle\phi| + |\phi\rangle\langle\psi|)\right]|\psi\rangle \\ &= \frac{H|\psi\rangle + \langle H\rangle|\psi\rangle - 2\langle H\rangle|\psi\rangle}{2(\Delta E)^2} - \lambda\tilde{H}|\psi\rangle - \langle\phi|\psi\rangle|\psi\rangle - |\phi\rangle \\ \Rightarrow \quad |\phi\rangle &= \left[\left(\frac{1}{2(\Delta E)^2} - \lambda\right)\tilde{H} + \langle\psi|\phi\rangle\right]|\psi\rangle\,, \end{aligned} \tag{4.61}$$

where we used Lemma 4.6. We insert the obtained expression for $|\phi\rangle$ back into (4.42) to obtain (after some working out)

$$\begin{aligned} 0 &= \frac{\{H,P\} - 2\langle H\rangle P}{2(\Delta E)^2} - \lambda\tilde{H} - P\left[\left(\frac{1}{2(\Delta E)^2} - \lambda\right)\tilde{H} - \langle\psi|\phi\rangle\right] \\ &\quad - \left[\left(\frac{1}{2(\Delta E)^2} - \lambda\right)\tilde{H} + \langle\psi|\phi\rangle\right]P \\ &= \frac{1}{2(\Delta E)^2}\left[HP + PH - 2\langle H\rangle P - P\tilde{H} - \tilde{H}P\right] + \lambda\left[-\tilde{H} + P\tilde{H} + \tilde{H}P\right] \\ &\overset{*}{=} \lambda\left[-\tilde{H} + P\tilde{H} + \tilde{H}P\right] \\ \Rightarrow \quad \tilde{H} &= P\tilde{H} + \tilde{H}P, \end{aligned} \tag{4.62}$$

where for $*$ we used $\tilde{H} = H - \langle H\rangle$ – the topic of Lemma 4.7 – and the last implication holds as $\lambda \neq 0$. Thus, the proposition is proven. □

[13] Multiply both sides by $\tilde{H}$ and use that the trace is invariant under cyclic permutations.

[14] Although this might seem like a cyclic argument, it is not as Proposition 4.9 does not use the contents of this lemma.

Proposition 4.10. *Equations* (4.41) *and* (4.42) *together imply that another equation to solve is given by*

$$\partial_t \tilde{H} \left|\psi\right\rangle = 0. \tag{4.63}$$

Proof. We can also plug equation (4.61) into (4.41), resulting in

$$\begin{aligned} 0 &= i\partial_t \left[\frac{H - \langle H\rangle}{2(\Delta E)^2}\right] - i\partial_t \left\{\left[\left(\frac{1}{2(\Delta E)^2} - \lambda\right)\tilde{H} + \langle\psi|\phi\rangle\right]|\psi\rangle\right\} \\ &\quad + H\left[\left(\frac{1}{2(\Delta E)^2} - \lambda\right)\tilde{H} + \langle\psi|\phi\rangle\right]|\psi\rangle \\ &= i\partial_t[\lambda\tilde{H}]\,|\psi\rangle + \tilde{H}\left(\frac{1}{2(\Delta E)^2}\right)[H - i\partial_t]\,|\psi\rangle \\ &= \partial_t[\lambda\tilde{H}]\,|\psi\rangle\,, \end{aligned} \tag{4.64}$$

where we used the product rule, the Schrödinger equation, $\partial_t[\langle\psi|\phi\rangle] = 0$ [15] and the fact that H and $\tilde{H}$ commute. To transform this into a condition purely on $\tilde{H}$, though, we require to say something about λ. Fortunately, though, we have that $\partial_t\lambda = 0$.

To see this, left-multiply equation (4.64) by $\langle\psi|\,H$, such that we obtain

$$\begin{aligned} 0 = \langle\psi|H\partial_t[\lambda\tilde{H}]|\psi\rangle &= \langle\tilde{H}\partial_t[\lambda\tilde{H}]\rangle + \langle H\rangle\,\langle\partial_t[\lambda\tilde{H}]\rangle \\ &= \lambda\underbrace{\langle\tilde{H}\partial_t[\tilde{H}]\rangle}_{(1)} + \lambda\underbrace{\langle H\rangle\,\langle\partial_t\tilde{H}\rangle}_{(2)} \\ &\quad + \partial_t\lambda\underbrace{\langle H\rangle\,\langle\tilde{H}\rangle}_{(3)} + \partial_t\lambda\underbrace{\langle\tilde{H}^2\rangle}_{(4)}. \end{aligned} \tag{4.65}$$

We proceed with all terms in (4.65) separately.

(1): By assurances of the author of [4], we have that this term disappears. In a written exchange, he made clear that this is due to $\langle\partial_t\tilde{H}^2\rangle = \partial_t\,\langle\tilde{H}^2\rangle = 0$, where the first equality comes from Ehrenfest's theorem. Although I could not directly make the connection, this is the implied reason that the term (1) is cancelled.

(2): Since $[H, \tilde{H}] = 0$, by Ehrenfest's theorem we have that $\langle\partial_t\tilde{H}\rangle = \partial_t\,\langle\tilde{H}\rangle$. Then, since $\langle\tilde{H}\rangle = 0$ by Lemma 4.7, we so have that $\langle\partial_t\tilde{H}\rangle = 0$, and the term (2) is cancelled.

(3): Again, by Lemma 4.7, $\langle\tilde{H}\rangle = 0$ so that the term (3) is cancelled.

(4): By the first equality in Lemma 4.8, we have that

$$\langle\tilde{H}^2\rangle = (\Delta E)^2 \implies \partial_t\lambda\,\langle\tilde{H}^2\rangle = \partial_t\lambda(\Delta E)^2. \tag{4.66}$$

Thus, equation (4.65) becomes

$$\partial_t\lambda(\Delta E)^2 = 0 \implies \partial_t\lambda = 0. \tag{4.67}$$

[15] This is justified, for $\partial_t[\langle\psi|\phi\rangle] = \langle\partial_t\psi|\phi\rangle + \langle\psi|\partial_t\phi\rangle = \langle\partial_t\psi|\phi\rangle - \langle\partial_t\psi|\phi\rangle = 0$.

Combining this fact and equation (4.64), we are thus justified to say

$$\partial_t \tilde{H} \left|\psi\right\rangle = 0, \tag{4.68}$$

showing the proposition. □

Although this is a nice equation, it is not as nice as it could be. Since we have now formulated our problem in terms of $\tilde{H}$, it is only natural to want a corresponding state with which it satisfies the Schrödinger equation – after all, the Schrödinger equation is the foremost tool of the quantum physicist. This is the topic of the following definition.

Definition 4.11. We define $\left|\tilde{\psi}\right\rangle := \exp\left(i\int_0^t \langle H\rangle \, dt'\right)\left|\psi\right\rangle$. In this way, the Schrödinger equation holds with $\tilde{H}$, i.e.

$$i\partial_t \left|\tilde{\psi}\right\rangle = \tilde{H}\left|\tilde{\psi}\right\rangle. \tag{4.69}$$

In particular, we also note that $\tilde{P} = \left|\tilde{\psi}\right\rangle\!\left\langle\tilde{\psi}\right| = P$, as the additional phases cancel.

Therefore, we rephrase equation (4.63) into

$$\partial_t \tilde{H} \left|\tilde{\psi}\right\rangle = 0. \tag{4.70}$$

This rewriting is allowed, as the difference between $\left|\psi\right\rangle$ and $\left|\tilde{\psi}\right\rangle$ is merely a phase factor, so that it cannot affect the total expression equaling to zero.

Thus, to recap this subsection, the equations we need to solve have become

$$\boxed{\tilde{H} = \tilde{H}P + P\tilde{H}} \tag{4.71}$$

and

$$\boxed{\partial_t \tilde{H} \left|\tilde{\psi}\right\rangle = 0.} \tag{4.72}$$

The solution to this system will be presented in the following subsection.

4.2.2 Solving The Equations

Now that we have simplified the equations, we are *almost* in a position to solve them. We shall have to do only a little more algebra, after which we find we are solving the physicists' favorite – a harmonic oscillator.

First of all, we would like to do something with equation (4.71). Notice that, since $\tilde{P} = P$, we have $\tilde{H}$ acting upon bra and ket $\tilde{\psi}$ states. And, by the very definition of $\left|\tilde{\psi}\right\rangle$, the Schrödinger equation applies. Therefore, we can rewrite (4.71) using the Schrödinger equation – namely to

$$\tilde{H} = i\left(\left|\partial_t\tilde{\psi}\right\rangle\!\left\langle\tilde{\psi}\right| - \left|\tilde{\psi}\right\rangle\!\left\langle\partial_t\tilde{\psi}\right|\right). \tag{4.73}$$

Regarding $\tilde{H}$, we also have the following lemma.

Lemma 4.12. *We have that $\tilde{H}$ is independent of time, i.e.*

$$\partial_t \tilde{H} = 0. \tag{4.74}$$

Proof. Let us carry out the derivation and see what we work with.

$$\partial_t \tilde{H} = i\left(|\partial_t^2\tilde{\psi}\rangle\langle\tilde{\psi}| + |\partial_t\tilde{\psi}\rangle\langle\partial_t\tilde{\psi}| - |\partial_t\tilde{\psi}\rangle\langle\partial_t\tilde{\psi}| - |\tilde{\psi}\rangle\langle\partial_t\tilde{\psi}|\right) = i\left(|\partial_t^2\tilde{\psi}\rangle\langle\tilde{\psi}| - |\tilde{\psi}\rangle\langle\partial_t^2\tilde{\psi}|\right). \tag{4.75}$$

Then, notice that we have

$$\tilde{H}^2 P = \tilde{H}^2 |\tilde{\psi}\rangle\langle\tilde{\psi}| = -|\partial_t^2\tilde{\psi}\rangle\langle\tilde{\psi}|, \tag{4.76}$$

and similarly $P\tilde{H}^2 = -|\tilde{\psi}\rangle\langle\partial_t^2\tilde{\psi}|$. Then, using equation (4.71), we can rewrite $\tilde{H}^2 P = P\tilde{H}^2 = \tilde{H}^2 - \tilde{H}P\tilde{H}$. It then becomes clear that

$$\partial_t \tilde{H} = i\left(P\tilde{H}^2 - \tilde{H}^2 P\right) = i\left(\tilde{H}^2 - \tilde{H}P\tilde{H} - \tilde{H}^2 + \tilde{H}P\tilde{H}\right) = 0, \tag{4.77}$$

so that the lemma is proven. □

This allows us to rewrite (4.73) slightly: since it does not change with time, $\tilde{H} = \tilde{H}(0)$. That is,

$$\tilde{H} = i\left(|\partial_t\tilde{\psi}(0)\rangle\langle\tilde{\psi}(0)| - |\tilde{\psi}(0)\rangle\langle\partial_t\tilde{\psi}(0)|\right) \tag{4.78}$$

Now, having found an expression for $\tilde{H}$ which does not depend on itself, we can have a go at solving equation (4.72). The following lemma aids in this, as it allows us to rewrite this equation into a more familiar form.

Lemma 4.13. *Using equation* (4.73) *as* $\tilde{H}$*,* (4.72) *is equivalent to*

$$(1 - \tilde{P})\,|\partial_t^2\tilde{\psi}\rangle = 0. \tag{4.79}$$

Proof. We start by taking the derivative of $\tilde{H}$ as in (4.73),[16] yielding

$$\begin{aligned}\partial_t \tilde{H} &= i\partial_t\left[|\partial_t\tilde{\psi}\rangle\langle\tilde{\psi}| - |\tilde{\psi}\rangle\langle\partial_t\tilde{\psi}|\right] \\ &= i\left(|\partial_t^2\tilde{\psi}\rangle\langle\tilde{\psi}| - |\tilde{\psi}\rangle\langle\partial_t^2\tilde{\psi}|\right).\end{aligned}$$

Therefore, by right-multiplying with $|\tilde{\psi}\rangle$ we get

$$\begin{aligned}0 = \partial_t\tilde{H}\,|\tilde{\psi}\rangle &= i\left(|\partial_t^2\tilde{\psi}\rangle\langle\tilde{\psi}| - |\tilde{\psi}\rangle\langle\partial_t^2\tilde{\psi}|\right)|\tilde{\psi}\rangle \\ &= i\left(|\partial_t^2\tilde{\psi}\rangle - |\tilde{\psi}\rangle\,\langle\partial_t^2\tilde{\psi}|\tilde{\psi}\rangle\right) \\ &= i\left(1 - |\tilde{\psi}\rangle\langle\tilde{\psi}|\right)|\partial_t^2\tilde{\psi}\rangle \\ \Rightarrow \quad &\left(1 - \tilde{P}\right)|\partial_t^2\tilde{\psi}\rangle = 0,\end{aligned}$$

showing the lemma. □

[16] This is slightly awkward due to to Lemma 4.12, but it is only the form we care about, not the actual value of the derivative.

Equation (4.79) is an equation that looks recognizable: it is *almost* an harmonic oscillator, owing to its double derivatives. The one requirement missing, though, is the constant – the "angular frequency." In the current form, there is no mention of it. However, the following lemma will introduce it.

Lemma 4.14. *We have that*

$$\tilde{P}\,|\partial_t^2\tilde{\psi}\rangle = -\omega^2\,|\tilde{\psi}\rangle\,, \tag{4.80}$$

where $\omega^2 := \mathrm{Tr}(\tilde{H})/2$.

Proof. Using the definition of $\tilde{P}$, we find

$$|\tilde{\psi}\rangle\langle\tilde{\psi}|\;|\partial_t^2\tilde{\psi}\rangle = \langle\tilde{\psi}|\partial_t^2\tilde{\psi}\rangle\;|\tilde{\psi}\rangle = -\langle\partial_t\tilde{\psi}|\partial_t\tilde{\psi}\rangle\;|\tilde{\psi}\rangle = -\langle\tilde{H}^2\rangle\;|\tilde{\psi}\rangle\,. \tag{4.81}$$

Since we know from Lemma 4.8 that $\langle\tilde{H}^2\rangle = \omega^2$, the lemma is proven. □

Hence, the equation to be solved now becomes simply

$$\left(\partial_t^2 + \omega^2\right)\;|\tilde{\psi}\rangle = 0, \tag{4.82}$$

which is precisely the harmonic oscillator we were on about earlier. This we can solve. We assume initial conditions $|\tilde{\psi}(0)\rangle$ and $|\partial_t\tilde{\psi}(0)\rangle$, and make an ansatz for our solution to be of the form

$$|\tilde{\psi}\rangle = A\cos(\omega t) + B\sin(\omega t), \tag{4.83}$$

where A and B belong to state space. Then, evaluating at $t = 0$ implies $A = |\tilde{\psi}(0)\rangle$, and evaluating the derivative at $t = 0$ implies $B = \frac{1}{\omega}\,|\partial_t\tilde{\psi}(0)\rangle$. Our expression for $|\tilde{\psi}\rangle$ so becomes

$$|\tilde{\psi}\rangle = \cos(\omega t)\,|\tilde{\psi}(0)\rangle + \frac{\sin(\omega t)}{\omega}\,|\partial_t\tilde{\psi}(0)\rangle\,. \tag{4.84}$$

We are now very close to the solution. What still rests us to do is to incorporate the initial and final states – $|\psi_i\rangle$ and $|\psi_f\rangle$ respectively – into this expression, as well as into the expression of the Hamiltonian (4.73). The initial state is easy: we simply set $|\tilde{\psi}(0)\rangle := |\psi_i\rangle$.

A word needs to be said regarding the final state, though: we must first orthonormalize the final state w.r.t. the initial state. Were we to not carry out this orthonormalization, and simply write $|\tilde{\psi}\rangle$ as a superposition of $|\psi_i\rangle$ and $|\psi_f\rangle$, then in guaranteeing that $|\tilde{\psi}\rangle$ is normalized – as physical states need to be – the inner product $\langle\psi_f|\psi_i\rangle$ would play a role. And since we *cannot* know the inner product (as the initial and final states are arbitrary), we need to anticipate and make sure that it no longer plays a role.

This is precisely achieved by orthonormalizing the final state w.r.t. the initial state. The relevant inner product will then be zero, which is what we require to show that $|\tilde{\psi}\rangle$ is normalized. The orthonormalization is carried out using the *Gram-Schmidt orthonormalization algorithm*, which we assume the reader is familiar with. This the topic of the following definition.

Definition 4.15. In agreeance with the Gram-Schmidt orthonormalization algorithm, the orthonormalized version of $|\psi_f\rangle$, $|\psi'_f\rangle$, is defined as follows:

$$|\psi'_f\rangle := \frac{|\psi_f\rangle - \langle\psi_f|\psi_i\rangle\,|\psi_i\rangle}{\sqrt{1-|\langle\psi_f|\psi_i\rangle|^2}}. \tag{4.85}$$

Lemma 4.16. *We have that*

$$|\psi'_f\rangle = \tfrac{1}{\omega}\,|\partial_t\tilde{\psi}(0)\rangle\,. \tag{4.86}$$

Proof. By construction of $|\psi'_f\rangle$, it holds that $\langle\psi_i|\psi'_f\rangle = 0$. Furthermore,

$$\langle\tilde{\psi}|\partial_t\tilde{\psi}\rangle = -i\,\langle\tilde{\psi}|\tilde{H}|\tilde{\psi}\rangle = \langle\tilde{H}\rangle = 0, \tag{4.87}$$

which holds for all t – in particular, it holds for $t = 0$. Evaluating at $t = 0$ gives us

$$\langle\tilde{\psi}(0)|\partial_t\tilde{\psi}(0)\rangle = \langle\psi_i|\partial_t\tilde{\psi}(0)\rangle\,, \tag{4.88}$$

by definition of $|\psi_i\rangle$. Thus, since $|\tilde{\psi}\rangle$ is spanned by $\{\,|\psi_i\rangle\,,\,|\psi'_f\rangle\}$, we must have that

$$|\psi'_f\rangle \propto |\partial_t\tilde{\psi}(0)\rangle\,. \tag{4.89}$$

For now, let us dub the proportion constant C, i.e. $|\psi'_f\rangle = C\,|\partial_t\tilde{\psi}(0)\rangle$. Then, substitution into equation (4.84) yields

$$|\tilde{\psi}\rangle = \cos(\omega t)\,|\psi_i\rangle + \frac{\sin(\omega t)}{C\omega}\,|\psi'_f\rangle\,. \tag{4.90}$$

Finally, utilizing $|\tilde{\psi}|^2 = 1$ for all t we get

$$1 = \cos^2(\omega t) + \frac{\sin^2(\omega t)}{C^2\omega^2} \implies C = \frac{1}{\omega}, \tag{4.91}$$

so that $|\psi'_f\rangle = \frac{1}{\omega}\,|\partial_t\tilde{\psi}(0)\rangle$ as desired. □

Using Lemma 4.16 and the definition of $|\psi_i\rangle$, we can rewrite equations (4.78) and (4.84) into the solutions of the unconstrained problem. We present

$$\boxed{\begin{aligned}|\tilde{\psi}\rangle &= \cos(\omega t)\,|\psi_i\rangle + \sin(\omega t)\,|\psi'_f\rangle \\ &= |\psi_i\rangle\left(\cos(\omega t) - \frac{\langle\psi_f|\psi_i\rangle\sin(\omega t)}{\sqrt{1-|\langle\psi_f|\psi_i\rangle|^2}}\right) + |\psi_f\rangle\left(\frac{\sin(\omega t)}{\sqrt{1-|\langle\psi_f|\psi_i\rangle|^2}}\right)\end{aligned}} \tag{4.92}$$

$$\boxed{\tilde{H} = i\omega\left(\,|\psi'_f\rangle\langle\psi_i| - |\psi_i\rangle\langle\psi'_f|\,\right)} \tag{4.93}$$

as the solution to the unconstrained quantum brachistochrone problem.

Remark 4.17. We have phrased our solution in terms of the traceless Hamiltonian $\tilde{H}$ only, for we solved the problem using *it* rather than the full Hamiltonian H. This effectively means that the unconstrained quantum brachistochrone problem is solved by any Hamiltonian H which has as its traceless part the determined $\tilde{H}$. A suggested form is $H(t) = \tilde{H} + \langle F(t) \rangle$, for

$$H(t) - \mathrm{Tr}(H(t))/n = \tilde{H} - \mathrm{Tr}(\tilde{H})/n + \langle F(t) \rangle - \mathrm{Tr}(\langle F(t) \rangle)/n = \tilde{H}, \tag{4.94}$$

for any appropirate F and $\tilde{H}$ as in (4.93).

In Figure 4.1, we have parametrized some solutions to the problem in t, with varying inner products $\langle \psi_f | \psi_i \rangle$. Note that in this modelling, we only considered real inner products: complex inner products would be harder to take into account. Ultimately, the shortest time in which a transformation of state can take place is the subject of the following proposition.

Proposition 4.18. *The shortest time in which a transition from $|\psi_i\rangle$ to $|\psi_f\rangle$ can take place is*

$$\boxed{T = \frac{1}{|\omega|} \arccos |\langle \psi_f | \psi_i \rangle|.} \tag{4.95}$$

Proof. We require that $|\tilde{\psi}(T)\rangle := |\psi_f\rangle$. Utilizing equation (4.92), then, we obtain

$$\langle \psi_f | \tilde{\psi}(T) \rangle = 1 = \cos(\omega T) \langle \psi_f | \psi_i \rangle + \sin(\omega T) \langle \psi_f | \psi_f' \rangle . \tag{4.96}$$

Or, using Definition 4.15,

$$1 = \cos(\omega T) \langle \psi_f | \psi_i \rangle + \sin(\omega T) \frac{1 - \langle \psi_f | \psi_i \rangle^2}{\sqrt{1 - |\langle \psi_f | \psi_i \rangle|^2}}. \tag{4.97}$$

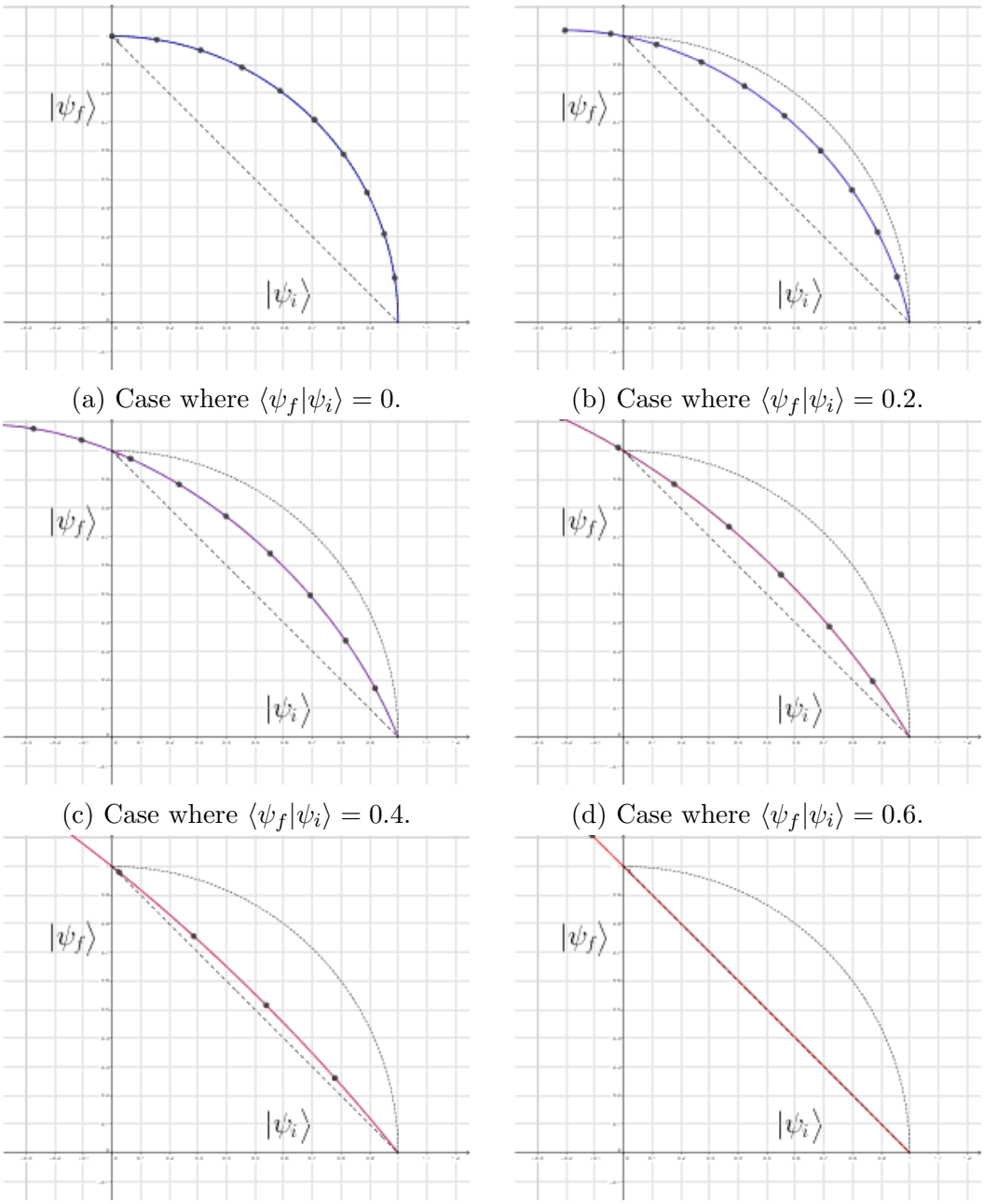
Evidently, $T = \arccos |\langle \psi_f | \psi_i \rangle| / |\omega|$ is a solution, using the conversion formula $\sin(A) = \sqrt{1 - \cos^2(A)}$. This finishes the proof.

Notice that taking the absolute value of $\langle \psi_f | \psi_i \rangle$ is necessary, as that ensures the square root resulting from the sine cancels against that from the substitution of $|\psi_f'\rangle$. □

Thus, from this point onward, the quantum brachistochrone problem can be considered to have been solved. Remember, though, that this solution only holds in an ideal world, as no constraints other than a trivial one have been imposed on the system. The next subsection will examine the case where additional constraints are imposed on the system, and the solution such a system has.

4.3 Case with constraints

The previous subsection discussed what happens when we assume there are no constraints on the Hamiltonian H besides the trivial one. However, this does not represent a physical system – in the real world, we have limitations, for instance on equipment. Thus, it is also imperative to see how these constraints play into the quantum brachistochrone problem.



(a) Case where $\langle\psi_f|\psi_i\rangle = 0$.

(b) Case where $\langle\psi_f|\psi_i\rangle = 0.2$.

(c) Case where $\langle\psi_f|\psi_i\rangle = 0.4$.

(d) Case where $\langle\psi_f|\psi_i\rangle = 0.6$.

(e) Case where $\langle\psi_f|\psi_i\rangle = 0.8$.

(f) Case where $\langle\psi_f|\psi_i\rangle = 0.99$.

Figure 4.1: The coefficients of $|\psi_i\rangle$ and $|\psi_f\rangle$ as a function of time in optimized transition, where different $\langle\psi_f|\psi_i\rangle$ are modelled. The points along the line represent timesteps of $\pi/20$ each, with $(1, 0)$ at $t = 0$. Note that the squares of the coefficients does not necessarily add up to 1: this is due to $\langle\psi_f|\psi_i\rangle \neq 0$ in general. Moreover, notice that with increasing $\langle\psi_f|\psi_i\rangle$, the time required to reach $|\psi_f\rangle$ decreases, as one would expect: at $\langle\psi_f|\psi_i\rangle = 0.8$, for instance, the time required is only approx. 40% of the orthogonal case.

For convenience, define

$$M := \sum_{j=1}^{m} \lambda_j \frac{\delta f_j}{\delta H}. \tag{4.98}$$

Thus, the equations which now need solving become

$$\boxed{i\partial_t \left[\frac{H - \langle H \rangle}{2(\Delta E)^2} \right] |\psi\rangle - i\partial_t |\phi\rangle + H |\phi\rangle = 0,} \tag{4.99}$$

and

$$\boxed{\frac{\{H, P\} - 2\langle H \rangle P}{2(\Delta E)^2} - M - (\, |\psi\rangle\langle\phi| \, + \, |\phi\rangle\langle\psi| \,) = 0.} \tag{4.100}$$

Fortunately, the way of solving these equations is quite similar to the previous subsection. Before we commence with that, though, we direct your attention to the following lemma.

Lemma 4.19. *We have that*

$$\langle\psi|\phi\rangle = -\langle\phi|\psi\rangle - \langle M \rangle . \tag{4.101}$$

Proof. We take the expectation of equation (4.100) to obtain

$$\left\langle \frac{\{H, P\} - 2\langle H \rangle P}{2(\Delta E)^2} \right\rangle - \langle M \rangle - \langle\phi|\psi\rangle - \langle\psi|\phi\rangle = 0. \tag{4.102}$$

Observing that

$$\langle HP \rangle = \langle\psi|H |\psi\rangle\langle\psi| \, |\psi\rangle = \langle H \rangle , \tag{4.103}$$

the same holding for $\langle PH \rangle$ and $\langle P \rangle = 1$ trivially, we have that the first term cancels. Thus, what is left is

$$-\langle M \rangle - \langle\phi|\psi\rangle - \langle\psi|\phi\rangle \implies \langle\psi|\phi\rangle = -\langle\phi|\psi\rangle - \langle M \rangle , \tag{4.104}$$

proving the lemma. □

Notice the similarity between Lemmas 4.19 and 4.6: both say something about the inner products of ψ and ϕ, though the route taken to show their respective properties diverge considerably.

With Lemma 4.19 in hand, we can proceed with solving our equations. We consider the following proposition first.

Proposition 4.20. *Equation* (4.100) *implies that*

$$\boxed{F = \{F, P\} = FP + PF,} \tag{4.105}$$

holds. Here, $F := M - \langle M \rangle P$.

Proof. The idea is similar to before. Let us first solve for $|\phi\rangle$ by multiplying (4.100) with $|\psi\rangle$, obtaining

$$\begin{aligned} 0 &= \left(\frac{\{H,P\} - 2\langle H\rangle P}{2(\Delta E)^2}\right)|\psi\rangle - M|\psi\rangle - |\psi\rangle\langle\phi|\,|\psi\rangle - |\phi\rangle\langle\psi|\,|\psi\rangle \\ &= \left(\frac{H-\langle H\rangle}{2(\Delta E)^2}\right)|\psi\rangle - M|\psi\rangle - \langle\phi|\psi\rangle\,|\psi\rangle - |\phi\rangle \\ &\implies |\phi\rangle = \left[\frac{H-\langle H\rangle}{2(\Delta E)^2} - M - \langle\phi|\psi\rangle\right]|\psi\rangle\,. \end{aligned} \tag{4.106}$$

A similar equation can be obtained for $\langle\phi|$. Then, we plug ϕ back into equation (4.100) and behold

$$\begin{aligned} 0 &= \frac{\{H,P\} - 2\langle H\rangle P}{2(\Delta E)^2} - M - P\left[\frac{H-\langle H\rangle}{2(\Delta E)^2} - M - \langle\psi|\phi\rangle\right] - \left[\frac{H-\langle H\rangle}{2(\Delta E)^2} - M - \langle\phi|\psi\rangle\right]P \\ &= \frac{1}{2(\Delta E)^2}\left(HP + PH - 2\langle H\rangle P - PH + \langle H\rangle - HP + \langle H\rangle P\right) \\ &\quad - M + PM + \langle\psi|\phi\rangle P + MP + \langle\phi|\psi\rangle \\ &\overset{*}{=} -M + PM + (-\langle\phi|\psi\rangle - \langle M\rangle)P + MP + \langle\phi|\psi\rangle P \\ &= -M + PM + MP - \langle M\rangle P, \end{aligned} \tag{4.107}$$

where for $*$ we used Lemma 4.19. Then, defining $F := M - \langle M\rangle P$, we see that

$$\begin{aligned} \{F,P\} - F &= (M - \langle M\rangle P)P + P(M - \langle M\rangle P) - (M - \langle M\rangle P) \\ &= MP + PM - M - \langle M\rangle P \\ &\overset{**}{=} 0, \end{aligned} \tag{4.108}$$

where $**$ holds as previously established. Therefore, we have that

$$F = \{F,P\}, \tag{4.109}$$

which finishes the proof. □

Before we go on to the next proposition, we offer the following lemma.

Lemma 4.21. *We have that M and H commute.*

Proof. It will be sufficient to show that $\delta f_j/\delta H$ commutes with H for all j; the lemma then readily follows.

Recall that $f_j : \mathcal{H} \to \mathbb{R}$, i.e. $f_j(H)$ is simply a scalar. Trivially, $Hf_j(H) = f_j(H)H$, which leads to

$$\frac{\delta(f_j H)}{\delta H} = \frac{\delta(H f_j)}{\delta H}. \tag{4.110}$$

By the product rule, we can expand both sides of this equation to find

$$\frac{\delta f_j}{\delta H}H + f_j = f_j + H\frac{\delta f_j}{\delta H} \implies H\frac{\delta f_j}{\delta H} = \frac{\delta f_j}{\delta H}H. \tag{4.111}$$

Since the value of j is arbitrary, it holds for all j. Therefore, the lemma is proven. □

Proposition 4.22. *Equations* (4.99) *and* (4.100) *together imply that*

$$\boxed{\left(\partial_t F + i[\tilde{H}, F]\right)|\psi\rangle = 0.} \tag{4.112}$$

This equation is also known as the quantum brachistochrone equation*, cf. [5].*

Proof. We take the obtained expression for $|\phi\rangle$ (4.106) and plug it in to (4.99) to obtain (after some working out and with aid of the Schrödinger equation)

$$\left(i\partial_t M + i\partial_t \langle\psi|\phi\rangle + MH + \langle\psi|\phi\rangle H - HM - H\langle\psi|\phi\rangle\right)|\psi\rangle = 0. \tag{4.113}$$

The fourth and sixth terms cancel against each other. Furthermore, the second term cancels as

$$\begin{aligned}\partial_t \langle\psi|\phi\rangle |\psi\rangle &= i\langle\psi|H|\phi\rangle - i\langle\psi|\left(H|\phi\rangle + \partial_t\left[\frac{H - \langle H\rangle}{2(\Delta E)^2}\right]|\psi\rangle\right) \\ &= i\frac{\langle\partial_t H - \partial_t\langle H\rangle\rangle}{2(\Delta E)^2} \\ &= 0,\end{aligned} \tag{4.114}$$

where we used (4.99) for $|\partial_t\phi\rangle$ and the last step holds as $\partial_t\langle H\rangle = \langle\partial_t H\rangle$ (Ehrenfest's theorem). Therefore, we have

$$\left(i\partial_t M + MH - HM\right)|\psi\rangle = 0. \tag{4.115}$$

An immediate consequence is $\partial_t M|\psi\rangle = 0$, as $[H, M] = 0$ as established in Lemma 4.21. In order to get an F into (4.115), we require to find $\partial_t(\langle M\rangle P)$; combining this with $\partial_t M$ we obtain $\partial_t F$. To that end, we calculate

$$\begin{aligned}\partial_t(\langle M\rangle P) = &\ \langle\partial_t\psi|M|\psi\rangle P + \langle\partial_t M\rangle P + \langle\psi|M|\partial_t\psi\rangle P + \langle M\rangle\partial_t P \\ &= \langle\partial_t M\rangle P - i\langle[M, H]\rangle P + \langle M\rangle\partial_t P \\ &= i\langle M\rangle\partial_t P \\ &= i\langle M\rangle(PH - HP) \\ &= i\langle M\rangle[P, H].\end{aligned} \tag{4.116}$$

Therefore, equation (4.115) can be rewritten to

$$\begin{aligned}&\left(\partial_t M - \partial_t(\langle M\rangle P) - i[M, H] + i\langle M\rangle[P, H]\right)|\psi\rangle = 0 \\ &\implies \left(\partial_t F + i[H, F]\right)|\psi\rangle = 0.\end{aligned} \tag{4.117}$$

□

The previous two propositions can be integrated into each other in order to form a solution for F.

Proposition 4.23. *Propositions 4.20 and 4.22 can be integrated in order to yield a solution*

$$\boxed{F = UF(0)U^\dagger,} \tag{4.118}$$

where $F(0)$ is the evaluation of F at $t = 0$ and satisfies $F(0) = \{F(0), P(0)\}$. U Is given as

$$\boxed{U(t) = \mathcal{T} \exp\left(-i \int_0^t \tilde{H}(t')\, dt'\right) \quad \Longleftrightarrow \quad \partial_t U = -i\tilde{H}(t)U(t) \text{ with } U(0) = I,} \tag{4.119}$$

where $\mathcal{T}$ is the time-ordered product. Notice that U is unitary, i.e. it satisfies $UU^\dagger = U^\dagger U = I$.

Proof. Instead of deriving this solution, we will only show that it works by demonstrating it satisfies the conditions on F as posed in Propositions 4.20 and 4.22.

To that end, let us first consider Proposition 4.22. We have that

$$\begin{aligned} \partial_t F = \partial_t[UF(0)U^\dagger] &= \partial_t U F(0) U^\dagger + UF(0)\partial_t U^\dagger \\ &= -i\tilde{H}UF(0)U^\dagger + iUF(0)U^\dagger\tilde{H} \\ &= -i\tilde{H}F + iF\tilde{H} \\ &= -i[\tilde{H}, F], \end{aligned} \tag{4.120}$$

so that therefore Proposition 4.22 holds.

Now consider Proposition 4.20. We have that

$$\begin{aligned} F = UF(0)U^\dagger &= U(F(0)P(0) + P(0)F(0))U^\dagger \\ &= UF(0)P(0)U^\dagger + UP(0)F(0)U^\dagger \\ &= FP + PF, \end{aligned} \tag{4.121}$$

so that it too is satisfied. □

Hence, with Proposition 4.23, the quantum brachistochrone problem for the constrained case is settled: there are abundant equations which can be used to fully solve a constrained quantum brachistochrone system.

In the next section, we shall apply the theory learned regarding constrained quantum brachistochrone problems in order to solve an example problem.

5 A Worked Example

In this section, we shall work an example utilizing the method we developed in the previous section. This way, through illustration, the method will become more tangible and, hopefully, will show how to use the derived formulas and such.

The setup is that we have a qubit – i.e. a two-state quantum particle – spin-1/2 particle under the influence of a magnetic field. The constraints we set for our system are

$$f_1(H) = \mathrm{Tr}(\tilde{H}^2)/2 - \omega^2 = 0 \tag{5.1}$$

and

$$f_2(H) = \mathrm{Tr}(\tilde{H}\sigma_z) = 0, \tag{5.2}$$

where σ_z is a *Pauli matrix*.

Although the constraints do not show it in their formulations, these actually have easy physical interpretations. The first constraint, as we have posed before, says that the energy involved only has a finite magnitude – basically, we do not allow infinite energies to come into play. The second constraint restricts our magnetic field to exist in merely two dimensions (canonically x and y). We will shortly formally prove the two implications mentioned here, specifically Lemma 5.4.

This section will carry on as follows. In the first section, we will define some key terms and lemmas we require in order to start working on the problem. Subsequently, we will find $U(t)$ and $\tilde{H}(t)$ – the most important operators we require in order to solve the given problem. Finally, by defining boundary conditions for our problem, we manage to deduce initial conditions and as such find a closed solution to our problem. Our last hurrah shall be to find the optimal time to make the given transition between the boundary states.

5.1 For starters

Before we begin with treating this problem, let us lay down some ground properties that we require. First off, we begin with the *Pauli matrices*. These will feature heavily in this example, and as such it is good to have a refresher as to their definition.

Definition 5.1. The *Pauli matrices* are the matrices σ_x, σ_y and σ_z, defined as

$$\sigma_x := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and } \sigma_z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.3}$$

Furthermore, we define $\vec{\sigma} := (\sigma_x, \sigma_y, \sigma_z)$, a vector consisting of 2×2 matrices. Some properties – which can be verified by inspection – are that they are Hermitian, traceless and satisfy $\sigma_j\sigma_k = \delta_{jk}I$.

The Pauli matrices are a well-known, special set of matrices: namely, they form a basis for the space of 2×2 traceless Hermitian matrices. This allows them to be extremely useful in describing the spin state of any spin-1/2 particle. For further information regarding these matrices, the reader can refer to any (introductory) textbook regarding quantum mechanics.

We now present two additional lemmas, which we need in order to get started wiht our example.

Lemma 5.2. *We have that* $\mathrm{Tr}(F) = 0$.

Proof. This holds as

$$\mathrm{Tr}(F) = \mathrm{Tr}(FP + PF) = 2\,\mathrm{Tr}(FP) = 2\,\langle F\rangle = 2\,\langle FP + PF\rangle = 4\,\langle F\rangle\,, \tag{5.4}$$

so that the only way the equalities hold is if $\langle F\rangle = 0$ and, by extension, $\mathrm{Tr}(F) = 0$. □

Lemma 5.3. *The Hamiltonian for our system (spin-1/2 qubit with a magnetic field) is given as*

$$H = \tilde{H} = -\vec{\sigma}\cdot\vec{B}, \tag{5.5}$$

where $\vec{B} = (B_x, B_y, B_z) \in \mathbb{R}^3$.

Proof. This Hamiltonian is based in the Pauli equation, which governs how spin-1/2 particles behave in the presence of electric and magnetic fields. Our version is obtained by considering a stationary particle, so that the kinetic part of the Pauli equation vanishes and the scalar potential becomes irrelevant. Then choosing constant appropriately yields the formula as posed in the lemma. □

With these three points in mind, we are ready to tackle the problem.

5.2 Finding U and $\tilde{H}$

In this section, we will do the brunt of the work: finding U and $\tilde{H}$. As can be seen from the formulae in the previous section, these are the workhorses of any solution we hope to find.

We employ the given constraints to deduce the following lemma.

Lemma 5.4. *We have that*

1. $B_z = 0$
2. $|\vec{B}| = |\omega|$

Proof. We shall first show 1. Owing to Lemma 5.3, we have that

$$\tilde{H} = -B_x\sigma_x - B_y\sigma_y - B_z\sigma_z \implies \tilde{H}\sigma_z = -B_z\sigma_z^2 = -B_z I, \tag{5.6}$$

where I is the identity matrix. Therefore, the only way to ensure $\mathrm{Tr}(\tilde{H}\sigma_z) = 0$ is to set $B_z = 0$.

Regarding 2, we notice that

$$\tilde{H}^2 = B_x^2\sigma_x^2 + B_y^2\sigma_y^2 + B_z^2\sigma_z^2 = (B_x^2 + B_y^2)I, \tag{5.7}$$

so that the trace is given by $\mathrm{Tr}(\tilde{H}) = 2(B_x^2 + B_y^2)$. Therefore, if $\tilde{H}$ is to satisfy the constraint (5.1), we must have that

$$B_x^2 + B_y^2 = \omega^2 \quad\iff\quad |\vec{B}| = |\omega|, \tag{5.8}$$

which finishes the proof. □

Furthermore, we have the following proposition.

Proposition 5.5. *We have that*

$$F = \lambda_1 \tilde{H} + \lambda_2 \sigma_z. \tag{5.9}$$

Proof. Since the variation of the first constraint yields $\tilde{H}$, and of the second σ_z, we have that

$$F = \lambda_1 \left(\tilde{H} - \langle \tilde{H} \rangle P \right) + \lambda_2 \left(\sigma_z - \langle \sigma_z \rangle P \right), \tag{5.10}$$

cf. the definition of F as in Proposition 4.20. However, since we know $\mathrm{Tr}(F) = 0$ (Lemma 5.2), we can thus say

$$\begin{aligned} 0 = \mathrm{Tr}(F) &= \mathrm{Tr}\left[\lambda_1 \left(\tilde{H} - \langle \tilde{H} \rangle P \right) + \lambda_2 \left(\sigma_z - \langle \sigma_z \rangle P \right) \right] \\ &\overset{*}{=} -\mathrm{Tr}\left[\left(\lambda_1 \langle \tilde{H} \rangle + \lambda_2 \langle \sigma_z \rangle \right) P \right] \\ &= -\lambda_1 \langle \tilde{H} \rangle - \lambda_2 \langle \sigma_z \rangle, \end{aligned} \tag{5.11}$$

where for $*$ we used that $\mathrm{Tr}(\tilde{H}) = \mathrm{Tr}(\sigma_z) = 0$. Therefore,

$$F = \lambda_1 \tilde{H} + \lambda_2 \sigma_z - \left(\lambda_1 \langle \tilde{H} \rangle + \lambda_2 \langle \sigma_z \rangle \right) P = \lambda_1 \tilde{H} + \lambda_2 \sigma_z, \tag{5.12}$$

proving the proposition. □

Then, we recall the solution for F to be

$$F = U F(0) U^\dagger, \tag{5.13}$$

as per Proposition 4.23. Furthermore, from the definition of U in Proposition 4.23, we recall that

$$\partial_t U = -i \tilde{H} U. \tag{5.14}$$

Thus, when right-multiplying equation (5.9) by U, we obtain

$$\partial_t U = \frac{i\lambda_2}{\lambda_1} \sigma_z U + U \frac{-i}{\lambda_1} F(0) = i\Omega \sigma_z U - iU(\tilde{H}(0) + \Omega \sigma_z), \tag{5.15}$$

where we defined $\Omega := \lambda_2/\lambda_1$, and used $F(0) = \lambda_1 \tilde{H}(0) + \lambda_2 \sigma_z$. Although this might look like a daunting differential equation to solve, the solution is comparatively simple:

$$U = \exp\left(i \int_0^t \Omega \sigma_z \, dt' \right) \exp\left(-i \int_0^t \tilde{H}(0) + \Omega \sigma_z \, dt' \right). \tag{5.16}$$

That the above expression satisfies the differential equation can be readily verified. Notice that we cannot yet remove the integral signs, as we do not know anything about the time evolution of the Lagrange multipliers λ_1 and λ_2. Fortunately, we have the following lemma to our rescue.

Lemma 5.6. *We have that λ_1 and λ_2 are constant.*

Proof. Let us explicitly write λ_1 and λ_2 as functions of time, i.e. $\lambda_1(t)$ and $\lambda_2(t)$. Then, following their use in F, we have that

$$F(t) = \lambda_1(t)\tilde{H}(t) + \lambda_2(t)\sigma_z. \tag{5.17}$$

However, recalling $F = UF(0)U^\dagger$, we rewrite so that

$$\lambda_1(t)\tilde{H}(t) + \lambda_2(t)\sigma_z = U(\lambda_1(0)\tilde{H}(0) + \lambda_2(0)\sigma_z)U^\dagger = \lambda_1(0)\tilde{H}(t) + \lambda_2(0)\sigma_z. \tag{5.18}$$

Thus, we conclude that $\lambda_1(t) = \lambda_1(0)$ and $\lambda_2(t) = \lambda_2(0)$ for all t, so that indeed they are constant. □

With Lemma 5.6 in hand, we can simplify equation (5.16) to read instead

$$\boxed{U = \exp\left(i\Omega\sigma_z t\right)\exp\left(-i\tilde{H}(0)t - i\Omega\sigma_z t\right).} \tag{5.19}$$

Note that we cannot contract the exponentials to simplify further. These are matrix exponentials, and so in order to contract, we would need to have that the exponents are commutative. Since this cannot be guaranteed in general, it is not done.

Then, it rests us to calculate $\tilde{H}$. This is done in the following proposition.

Proposition 5.7. *We have that $\tilde{H}$ is given as follows:*

$$\boxed{\tilde{H} = \exp\left(i\Omega\sigma_z t\right)\tilde{H}(0)\exp\left(-i\Omega\sigma_z t\right).} \tag{5.20}$$

Proof. Earlier, we established that $\partial_t U = -i\tilde{H}U$. Right-multiply with $U^\dagger$ and we obtain $\tilde{H} = i\partial_t U U^\dagger$. So far the theory, now the practice.

First, we ascertain anew $\partial_t U$ from equation (5.15) and slightly rewrite it:

$$\begin{aligned}
\partial_t U &= i\Omega\sigma_z U - iU(\tilde{H}(0) + \Omega\sigma_z) \\
&= i\Omega\sigma_z U - i\exp\left(i\Omega\sigma_z t\right)(\tilde{H}(0) + \Omega\sigma_z)\exp\left(-i\tilde{H}(0)t - i\Omega\sigma_z t\right) \\
&= i\Omega\sigma_z U - i\Omega\sigma_z U - i\exp\left(i\Omega\sigma_z t\right)\tilde{H}(0)\exp\left(-i\tilde{H}(0)t - i\Omega\sigma_z t\right) \\
&= -i\exp\left(i\Omega\sigma_z t\right)\tilde{H}(0)\exp\left(-i\tilde{H}(0)t - i\Omega\sigma_z t\right),
\end{aligned} \tag{5.21}$$

where for the second and third equalities we used that $A\exp(At) = \exp(At)A$. Then, determining $U^\dagger$ to be

$$U^\dagger = \exp\left(i\tilde{H}(0)t + i\Omega\sigma_z t\right)\exp\left(-i\Omega\sigma_z t\right), \tag{5.22}$$

we right-multiply it with $i\partial_t U$ to find

$$\begin{aligned}
\tilde{H} = i\partial_t U U^\dagger &= \exp\left(i\Omega\sigma_z t\right)\tilde{H}(0)\exp\left(-i\tilde{H}(0)t - i\Omega\sigma_z t\right) \\
&\quad\times\exp\left(i\tilde{H}(0)t + i\Omega\sigma_z t\right)\exp\left(-i\Omega\sigma_z t\right) \\
&= \exp(i\Omega\sigma_z t)\tilde{H}(0)\exp(-i\Omega\sigma_z t),
\end{aligned} \tag{5.23}$$

as the second and third exponentials cancel to form I. Hence, the proposition is proven. □

Having now found expressions for U and $\tilde{H}$, the rest of solving the problem is nearly trivial. In the following subsection, we will add initial and final conditions in order to further our example.

5.3 Boundary Conditions and The Solution

In the previous subsection we found the expressions for U and $\tilde{H}$ pertaining to our specific constraints. However, we cannot fully complete the problem without also imposing some initial and final values that our system should take.

In that spirit, we will set them. Our conditions are

$$P(0) = \frac{1}{2}(I + \sigma_x) = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad P(T) = \frac{1}{2}(I - \sigma_x) = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \tag{5.24}$$

Notice in particular that these P matrices imply the corresponding states as well. It can be shown trivially that

$$|\psi(0)\rangle = |\psi_i\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ and } |\psi(T)\rangle = |\psi_f\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{5.25}$$

by computing the outer products.

Then, we first pose the following lemma concerning $\tilde{H}(0)$.

Lemma 5.8. *We have that*

$$\boxed{\tilde{H}(0) = \omega\sigma_y.} \tag{5.26}$$

Proof. Consider $F(0)$ as

$$F(0) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{5.27}$$

Then, it should also hold that

$$\begin{aligned} F(0) = F(0)P(0) + P(0)F(0) &= \frac{1}{2}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 2a + b + c & 2b + a + d \\ 2c + a + d & 2d + b + c \end{pmatrix}, \end{aligned} \tag{5.28}$$

from which we infer that $a + d = b + c = 0$.

Now recall that the form of $F(0)$ is given by $F(0) \propto \tilde{H}(0) + \Omega\sigma_z$, and that

$$\tilde{H}(0) = B_x(0)\sigma_x + B_y(0)\sigma_y = \begin{pmatrix} 0 & B_x(0) + iB_y(0) \\ B_x(0) - iB_y(0) & 0 \end{pmatrix}. \tag{5.29}$$

Therefore, recalling that σ_z only has only diagonal elements, we must have that

$$b + c = 0 \implies (B_x(0) + iB_y(0)) + (B_x(0) - iB_y(0)) = 0 \implies B_x(0) = 0. \tag{5.30}$$

Then, by Lemma 5.4, we have that $B_y(0) = \pm\omega$, of which we choose the positive variant as the sign is irrelevant anyway. Hence,

$$\tilde{H}(0) = \begin{pmatrix} 0 & i\omega \\ -i\omega & 0 \end{pmatrix} = \omega\sigma_y, \tag{5.31}$$

which proves the lemma. □

Having now found the initial Hamiltonian, the rest of the problem can be "simply" solved by applying equations (5.19) and (5.20) to this initial Hamiltonian. This is the topic of the following Theorem, where we get the solution of this example.

Theorem 5.9. *We have that the example in this section is solved by*

1. $$\vec{B} = \begin{pmatrix} \sin(2\Omega t) & \cos(2\Omega t) & 0 \end{pmatrix}$$

2. $$|\psi(t)\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} \exp(i\Omega t)\left(\cos(\Omega' t) + \frac{\omega - i\Omega}{\Omega'}\sin(\Omega' t)\right) \\ \exp(-i\Omega t)\left(\cos(\Omega' t) + \frac{i\Omega - \omega}{\Omega'}\sin(\Omega' t)\right) \end{pmatrix}$$

3. $$\langle \vec{\sigma} \rangle (t) = \begin{pmatrix} \langle \sigma_x \rangle (t) \\ \langle \sigma_y \rangle (t) \\ \langle \sigma_z \rangle (t) \end{pmatrix} = \begin{pmatrix} \cos(2\Omega t)\cos(2\Omega' t) + \frac{\Omega}{\Omega'}\sin(2\Omega t)\sin(2\Omega' t) \\ -\sin(2\Omega t)\cos(\Omega' t) + \frac{\Omega}{\Omega'}\cos(2\Omega t)\sin(2\Omega' t) \\ \frac{\omega}{\Omega}\sin(2\Omega' t) \end{pmatrix}$$

Here, $\Omega' := \sqrt{\Omega^2 + \omega^2}$.

Proof. We will only give sketches of the proofs. For full calculations, we refer to [8].

1. The matrix exponentials $\exp(\pm i\Omega\sigma_z t)$ can be readily calculated, as σ_z is a diagonal matrix. Then, we utilize equation (5.20) to get

$$\begin{aligned} \tilde{H}(t) &= \exp(i\Omega\sigma_z t)\tilde{H}(0)\exp(-i\Omega\sigma_z t) \\ &= \begin{pmatrix} 0 & i\omega\exp(2i\Omega t) \\ -i\omega\exp(-2i\Omega t) & 0 \end{pmatrix} \\ &= -\omega\sin(2\Omega t)\sigma_x + \cos(2\Omega t)\sigma_y). \end{aligned} \tag{5.32}$$

   Then, recalling that $\tilde{H} = -\vec{\sigma}\cdot\vec{B}$, we have that

$$\vec{B} = \omega\begin{pmatrix} \sin(2\Omega t) & \cos(2\Omega t) & 0 \end{pmatrix}. \tag{5.33}$$

   Thus, the first claim is proven.

2. We can find $|\psi(t)\rangle$ by utilizing the formula

$$|\psi(t)\rangle = U|\psi(0)\rangle, \tag{5.34}$$

   where U is – as we solved – given by equation (5.19). Then, using a nifty formula for the matrix exponential,[17] U can be found using $\tilde{H}(0)$. The multiplication is then straightforward, so that $|\psi(t)\rangle$ can be found.

[17] The formula is

$$\exp\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{\Delta}\begin{pmatrix} \exp(\frac{a+d}{2})[\Delta\cosh(\frac{\Delta}{2}) + (a-d)\sinh(\frac{\Delta}{2})] & 2b\exp(\frac{a+d}{2})\sinh(\frac{\Delta}{2}) \\ 2c\exp(\frac{a+d}{2})\sinh(\frac{\Delta}{2}) & \exp(\frac{a+d}{2})[\Delta\cosh(\frac{\Delta}{2}) + (d-a)\sinh(\frac{\Delta}{2})] \end{pmatrix},$$

where $\Delta := \sqrt{(a-d)^2 + 4bc}$.

3. Since we know $|\psi(t)\rangle$, $\langle\vec{\sigma}\rangle(t) = \begin{pmatrix}\langle\sigma_x\rangle & \langle\sigma_y\rangle & \langle\sigma_z\rangle\end{pmatrix}(t)$ can be calculated by simply finding $\langle\psi(t)|\sigma_j|\psi(t)\rangle$ for $j = x, y, z$.

□

The previous theorem solves the example in closed form: we know everything we need in order to calculate any other quantities. In particular, we can calculate the optimal time duration needed to make the transition $|\psi_i\rangle \to |\psi_f\rangle$. This is the topic of the following proposition.

Proposition 5.10. *We have that the following three statements hold.*

1. *The possible time durations are given by*

$$|\omega|T = \frac{\pi}{2}\sqrt{\ell^2 - k^2}, \tag{5.35}$$

   where $\ell > k \geq 0$ and $k + \ell$ is odd.

2. *Evidently, the* optimal *time duration corresponds to $(\ell, k) = (1, 0)$, so that $|\omega|T = \pi/2$.*

3. *The possible frequencies Ω are given by*

$$\left|\frac{\Omega}{\omega}\right| = \frac{|k|}{\sqrt{\ell^2 - k^2}}. \tag{5.36}$$

   The same conditions apply to ℓ and k as in 1.

Proof. We shall mainly focus on point 1, as 2. and 3. follow naturally.

1. First off, notice that $|\psi(T)\rangle = |\psi_f\rangle \Rightarrow \langle\psi(T)|\vec{\sigma}|\psi(T)\rangle = \begin{pmatrix}-1 & 0 & 0\end{pmatrix}$. Immediately, we notice that $2\Omega' t = \ell\pi$ for some $\ell \in \mathbb{Z}$, so as to ensure $\langle\sigma_z\rangle = 0$. Also, we deduce that $2\Omega t = k\pi$ for some $k \in \mathbb{Z}$ so that, with ℓ in mind, $\langle\sigma_y\rangle = 0$ holds as well.

   As a result, $\langle\sigma_x\rangle = \cos(2\Omega t)\cos(\Omega' t) = \cos(\ell\pi)\cos(k\pi) = (-1)^{\ell+k}$. Thus, to ensure negative unity, we must have that $\ell + k$ is odd.

   Then, since $\Omega'^2 > \Omega^2$ (as $\omega \neq 0$), we must have that $\ell > k$, where strict inequality holds as we still need to ensure $\ell + k$ is odd. If strict inequality did not hold, one could choose $\ell = k$, implying that $\ell + k$ is divisble by 2.

   Using the definition of Ω', we determine $\omega^2 = \Omega'^2 - \Omega^2$, leading to

$$\omega^2 = \frac{1}{4T^2}\left((2\Omega' T)^2 - (2\Omega T)^2\right) = \frac{\pi^2}{4T^2}\left(\ell^2 - k^2\right). \tag{5.37}$$

   This implies

$$|\omega|T = \frac{\pi}{2}\sqrt{\ell^2 - k^2}, \tag{5.38}$$

   so that the formula in point 1. is shown.

2. The consequences for the optimal time duration are evident, so that point 2. is shown.

3. We can combine the formula derived under 1. together with $2\Omega T = k\pi$ to obtain
$$2\Omega\left(\frac{\pi}{2|\omega|}\sqrt{\ell^2-k^2}\right) = k\pi \implies \left|\frac{\Omega}{\omega}\right| = \frac{|k|}{\sqrt{\ell^2-k^2}}. \tag{5.39}$$

All three points have been proven, so that the proposition is shown. □

There are two things to remark regarding this proposition. One neat feature is that it precisely matches the optimal time given in Proposition 4.18 for our particular choice of $|\psi_i\rangle$ and $|\psi_f\rangle$. Notable is that this occurs despite, in our example, there being an additional constraint imposed. It thus seems that the additional constraint (5.2) did not impact *at all* the system's ability to transition states in the globally optimal time.

Secondly, the expressions under 1. and 3. of the proposition can be combined in order to yield the simple-looking formula
$$T = \frac{|k|\pi}{2\Omega}. \tag{5.40}$$

Therefore, the times in which the transition can take place – and hence the optimal time – turn out to not depend on the strength of the magnetic field ω, instead only on the frequency Ω. As expected, the relation is inverse proportional: the larger Ω is, the smaller the transition time T is allowed to be. In general, though, it would still be advisable to choose an ω of reasonable strength. This so that the expression under 1. allows T to be as small as possible still.

5.4 Physical interpretation: Bloch sphere

One option for a physical interpretation of the solution is as its projection on the *Bloch sphere*. The Bloch sphere is a visualization aid commonly used in quantum mechanics, as it allows for a tangible link to the classically four-dimensional state (two imaginary components). This is done by exploiting the fact that normalization must hold, so that, effectively, one of the four coordinates becomes redundant, and three remain. By constructing these three coordinates – known as the *Bloch vector* – in a clever manner, normalization must also hold for them. Hence, in sum, every state can be described by a point on a 2-dimensional sphere.

There are items that we will skip over in this treatise of the Bloch sphere (for example showing that indeed the Bloch vector is normalized) for the sake of brevity. For full details, we refer the reader to [7].

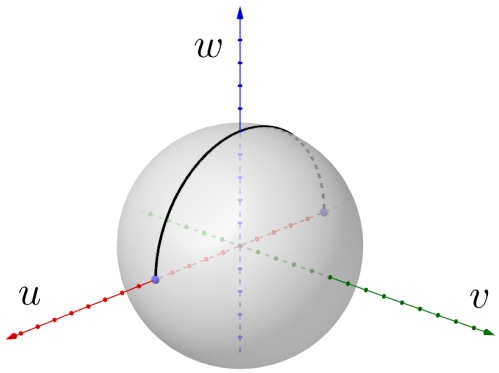
The Bloch vector is classically denoted by (u, v, w), where the quantities are given by $u = \rho_{01} + \rho_{10}$, $v = i(\rho_{01} - \rho_{10})$ and $w = \rho_{00} - \rho_{11}$, respectively. Here, ρ_{ij} indicates the ijth coordinate of the density matrix $\rho = |\psi(t)\rangle\langle\psi(t)|$. And so, since we know the state, the Bloch vector can be readily determined.

Omitting the calculations – as they are only calculations – we pose the Bloch vector associated with $|\psi(t)\rangle$ to be
$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \frac{\Omega}{\Omega'}\sin(2\Omega t)\sin(2\Omega' t) + \cos(2\Omega t)\cos(2\Omega' t) \\ \frac{\Omega}{\Omega'}\sin(2\Omega' t)\cos(2\Omega t) - \sin(2\Omega t)\cos(2\Omega' t) \\ \frac{\omega}{\Omega'}\sin(2\Omega' t) \end{pmatrix}. \tag{5.41}$$

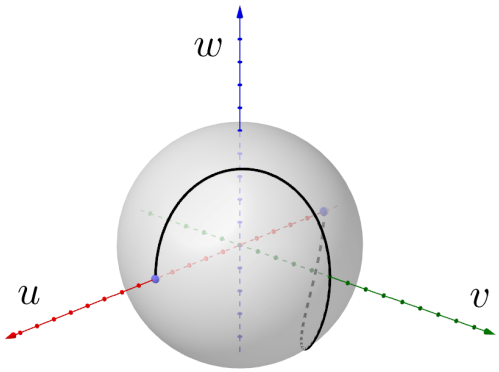
ℓ	k	Ω	$T\,(\times\pi/2)$
1	0	0	1
2	1	$1/\sqrt{3}$	$\sqrt{3}$
3	2	$2/\sqrt{5}$	$\sqrt{5}$
4	3	$3/\sqrt{7}$	$\sqrt{7}$
3	0	0	3
5	4	4/3	3
6	5	$5/\sqrt{11}$	$\sqrt{11}$

Table 1: Possible values for (ℓ, k) and their resulting Ω and T (we set $\omega \equiv 1$). The numbers have been ranked in ascending order of T.

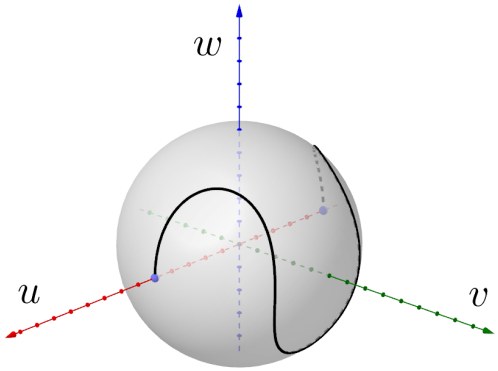
Then, all that we need before we can draw the Bloch sphere and associated vectors are some numbers to work with. For simplicity, we set $\omega \equiv 1$. Then, by choosing the constants ℓ and k, we can fully determine the system, i.e. the values of Ω, and T. A list of possible choices is given in Table 1. We can then model these on the Bloch sphere, which has been done in Figure 5.1 for the values given in Table 1. As one can see, the globally optimal time is indeed where we choose $(\ell, k) = (1, 0)$.



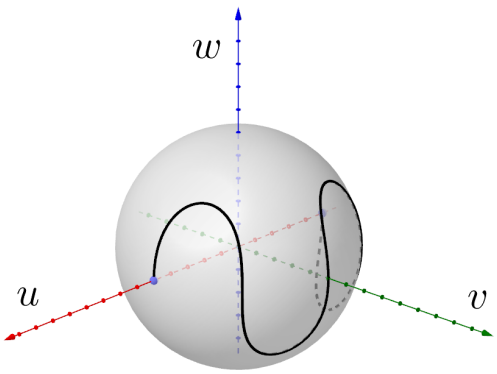
(a) Trajectory for $(\ell, k) = (1, 0)$.



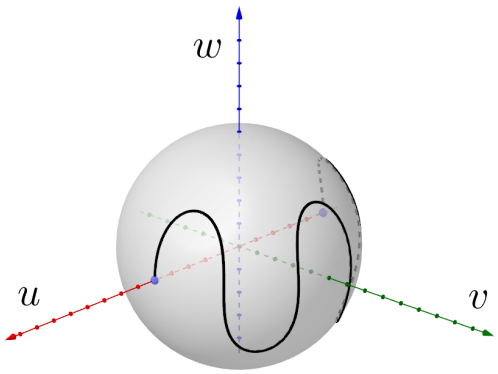
(b) Trajectory for $(\ell, k) = (2, 1)$.



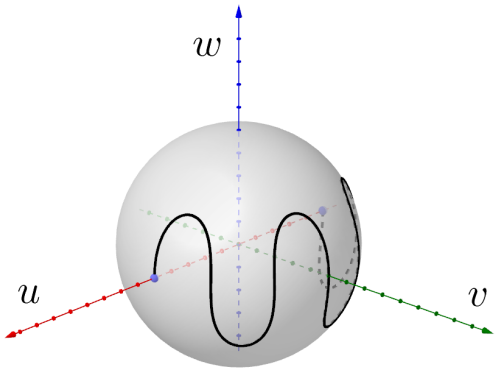
(c) Trajectory for $(\ell, k) = (3, 2)$.



(d) Trajectory for $(\ell, k) = (4, 3)$.



(e) Trajectory for $(\ell, k) = (5, 4)$.



(f) Trajectory for $(\ell, k) = (6, 5)$.

Figure 5.1: Some examples of trajectories on the Bloch sphere for the optimal state transitions belonging to the example treated in this section. The starting point is $(u, v, w) = (1, 0, 0)$, and the end point $(-1, 0, 0)$. Notice how the solutions have $0, 1, 2, \ldots$ nodes with the u-v plane for (a), (b), (c), etc. respectively. Moreover, notice how all save $(\ell, k) = (1, 0)$ do *not* define geodesics on the Bloch sphere despite being optimal solutions.

6 Quantum Computers: An Application

One might be surprised to find quantum computers discussed in a thesis which has a (supposedly) markedly different topic. However, the time-optimization of certain state transitions – the objective of the quantum brachistochrone problem – shares considerable similarities with quantum computation. In computation, the challenge has always been to do it faster and more efficiently. And considering that quantum computation is founded in the transitions of qubits between different states, optimizing these transitions thus yields faster computers.

The purpose of this section is thus to exemplify how the quantum brachistochrone might be used in physical applications, instead of remaining in the theorists' realm. We shall do this by showing how the quantum brachistochrone method can optimize unitary transformations between quantum states, the workhorse of any quantum computation.

The first section is devoted to giving a brief introduction to quantum computing. We shall explain the concept of a qubit and attempt to get a feel for them. Subsequently, we show that unitary transformations are to quantum computation as Boolean gates are to classical computation, showing their necessity.

Then, we enter the heart of this section. Namely, here we show how the aforementioned unitary transformations can be optimized using the methods of the quantum brachistochrone. We shall obtain similar results as we did in Section 4, except that they will be centered on the operator rather than the state.

Finally, we will apply the theory learned in the previous subsection to work the example of a specific unitary transformation.

Before we continue with the material, there are two things that need to be mentioned. First of all, provided here is a not a complete introduction to the subject of quantum computation, and that is not its purpose. We only need to be on the same page regarding the workings of qubits, which is what the first section accommodates. Secondly, precise proofs are omitted in this section for the sake of brevity and not losing sight of the goal – being convincing the reader of the practical applicability of the quantum brachistochrone. Elaborate proofs would eclipse this ultimate purpose.

6.1 Basics of Quantum Computing

6.1.1 Qubits

Just as a classical computer is at its heart served by bits, so a quantum computer is served by so-called qubits. These qubits are what make quantum computers so powerful, and namely for the following reason. Where bits can only attain two values (being 0 and 1), qubits are a *superposition* of the *states* $|0\rangle$ and $|1\rangle$. That is, any one qubit $|\psi\rangle$ can be described by

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle\,, \tag{6.1}$$

where a and b i) make sure that the state is normalized, i.e. $|a|^2 + |b|^2 = 1$, but also ii) indicate which way the state is "leaning." That is, if $a \gg b$, then, when measured, the state will more likely collapse to $|0\rangle$ than $|1\rangle$.

For a physical example of a qubit, this can be found in the spin of, say, an electron. Namely, spin is quantum property of the electron and can attain precisely two values – up or down – corresponding to the $|0\rangle$ and $|1\rangle$ states.

However, having only one qubit at our disposal is not much to look at. Supposing that our system is comprised of two qubits now, the superposition of *both qubits together* can be described as

$$|\psi\rangle = a\,|0\rangle_1\,|0\rangle_2 + b\,|0\rangle_1\,|1\rangle_2 + c\,|1\rangle_1\,|0\rangle_2 + d\,|1\rangle_1\,|1\rangle_2\,, \tag{6.2}$$

where a, b, c and d serve the same purpose as in the one qubit case. The joined kets indicate the base state of the respective qubit. Since the notation we defaulted to – subscripts indicating which qubit we are referring to – can be quite a hassle to write, we introduce the notation commonly employed in the field of quantum computation.

Notation 6.1. Instead of writing subscripts to indicate which qubit we are referring to, we join all qubits in the same ket and let the position in said ket indicate the qubit. That is,

$$|00\rangle \equiv |0\rangle_1\,|0\rangle_2\,. \tag{6.3}$$

Or, more generally for $\{a_i\}_{i=1}^n$, $a_i \in \{0,1\}$,

$$|a_1 a_2 ... a_n\rangle \equiv |a_1\rangle_1\,|a_2\rangle_2\,...\,|a_n\rangle_n\,. \tag{6.4}$$

In this notation, the qubit in equation (6.2) is expressed as

$$|\psi\rangle = a\,|00\rangle + b\,|01\rangle + c\,|10\rangle + d\,|11\rangle\,. \tag{6.5}$$

Some authors even take the notation a step further by taking the arguments of the kets as binary notation for numbers. For instance, since in binary $10 \equiv 2$, they would write $|2\rangle$ for $|10\rangle$.

6.1.2 Unitary transformations

So far, we have seen that qubits are the quantum analogy of regular bits, and seen how to express these qubits in notation. However, writing down the qubits is one thing; now, we want to do something with them. After all, computers rely on rapidly changing the states of bits – be it regular or quantum – to do calculations.

In a classical computer this rapid changing of states is done by using the well-known *Boolean gates*. Famous examples include the NOT-, XOR- and AND-gates. Essentially, a classical computer is a large array of Boolean gates strung together in a particualr way such that calculations can be performed.

For qubits, a very similar methodology can be applied – we also need some type of gate to manipulate qubits and make them perform calculations. The difference, however, is in the type of gate: whereas classical computers use Boolean gates, quantum computers use *quantum gates*, canonically designated as U. Not every operation one could conceivably subject a qubit to can be classified as a quantum gate, though. The properties a quantum gate U has to fulfill stand very close to the postulates of quantum mechanics. For our purposes, though, we can formulate the restrictions as three points, which are given below.

1. The operation U should be linear, i.e. we must ensure that

$$U \sum_{j=1}^{n} a_j \left|\psi_j\right\rangle = \sum_{j=1}^{n} a_j U \left|\psi_j\right\rangle \tag{6.6}$$

for given constants a_j and qubit states $\left|\psi_j\right\rangle$. However, this is almost a given as, generally, quantum operations are already assumed to be linear.

2. The operation U has to be length preserving. That is, if we apply a transformation $\left|\psi\right\rangle \to U\left|\psi\right\rangle$, then $|\psi|^2 = |U\psi|^2$. Under the additional assumption of normalization, this implies that, after U has been applied, we still are left with a normalized state.

3. The inner product of two qubit states has to be preserved under application of U. In other words, we must ensure that if $\left(\left|\phi\right\rangle, \left|\psi\right\rangle\right) \to \left(U\left|\phi\right\rangle, U\left|\psi\right\rangle\right)$, then $\langle\phi|\psi\rangle = \langle\phi|U^\dagger U|\psi\rangle$.

The three conditions above imply that the quantum gates are actually *unitary matrices*. For a justification, let us treat the conditions point for point. The first condition implies that U should be a matrix, due to the linearity; the second implies that U has (negative) unity determinant; and the third implies that $U^\dagger U = I \Rightarrow U^\dagger = U^{-1}$. These are precisely the properties of a unitary matrix, such it and the quantum gate that are identical.

An example of a unitary transformation is the quantum XOR-gate, which truth table is given in Table 2. The more natural way to write this down, though, is by

Input	Output
$\left\|00\right\rangle$	$\left\|00\right\rangle$
$\left\|01\right\rangle$	$\left\|01\right\rangle$
$\left\|10\right\rangle$	$\left\|11\right\rangle$
$\left\|11\right\rangle$	$\left\|10\right\rangle$

Table 2: Truth table for U_{XOR}, the quantum version of the XOR-gate

exploiting matrix notation in the canonical qubit basis $\{\left|00\right\rangle, \left|01\right\rangle, \left|10\right\rangle, \left|11\right\rangle\}$:

$$U_{\text{XOR}} : \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} a \\ b \\ d \\ c \end{pmatrix} \quad \Longleftrightarrow \quad U_{\text{XOR}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{6.7}$$

Further reference

Quantum computing is a vastly greater field than the two or so previous pages have explained. In order to get a fuller overview of the topic, the reader is recommended to explore introductory sources on the matter, for instance [9]

6.2 Optimized Unitary Transformations

As we have seen in the previous subsection, unitary transformations are vital to the existence of quantum computers. In fact, the speed at which these unitary operations can be applied to qubit states directly influences the speed at which your quantum computer performs, as you might expect. It is therefore only natural to attempt to optimize the time in which these unitary transformations take place, and investigate under which conditions optimization is ensured.

This is where the material discussed in this thesis comes in – specifically subsection 4.3. Namely, in this subsection we derived the solution for the generalized/constrained quantum brachistochrone problem, importantly using unitary matrices. With only a minor modification, the method derived for finding the time-optimal Hamiltonian can be altered to find the time-optimal unitary transformation.

In their paper on this very subject [5], Carlini *et al.* propose a five-step plan in order to find the time-optimized unitary transformation corresponding to a quantum gate f.

Plan 6.2. In order to find the time-optimal unitary transformation U_f corresponding to some function f (e.g. a quantum gate), the five steps below need to be followed.

1. Specify the constraints g_j such that $g_j(H) = 0$ for $j = 1, 2, \ldots, m$.
2. Write down the quantum brachistochrone equation, i.e.
$$i\partial_t F = [H, F], \tag{6.8}$$
where $F = \sum_{j=1}^{m} \lambda_j (\delta g_j(H)/\delta H)$. Note that this equation is the topic of Proposition 4.22.
3. Solve this quantum brachistochrone equation to obtain the Hamiltonian $H(t)$.
4. Integrate the equation
$$U = \mathcal{T} \exp\left(i \int_0^t H(t')\, dt' \right) \tag{6.9}$$
together with $U(0) = I$ to obtain $U(t)$. Note that this is the topic of Proposition 4.23.
5. Fix any constants still remaining by demanding that
$$U(T) = e^{i\chi} U_f, \tag{6.10}$$
where T is the end time.

Additionally, since $U(t)$ was the time-optimal way to carry out U_f, it follows that T is the optimal time in which the operation could be completed.

Their derivation in this paper follows, in broad lines, the same derivation published in [4]. However, one major difference is that here they refrain from ever mentioning the state, conforming to the *Heisenberg picture of quantum mechanics* in which only the operators vary with time, not the states. This last concept we also touched upon in subsection 4.3.

6.3 An example

Let us now treat a concrete example. We consider a two-qubit system, where we allow these qubits interact with each other, and we impose a magnetic field in the x- and y-directions. Then, we wish to implement the *SWAP-gate*, i.e. the transformation

$$U_{\text{SWAP}} : \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} a \\ c \\ b \\ d \end{pmatrix} \quad \Longleftrightarrow \quad U_{\text{SWAP}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{6.11}$$

in the time-optimal manner; that is, we want to work through Plan 6.2 for $U_f = U_{\text{SWAP}}$. Notice that the physical interpretation of U_{SWAP} is simply to swap the amplitudes belonging to the states $|01\rangle$ and $|10\rangle$.

In our case, we shall only briefly sketch steps one through four of Plan 6.2, and then treat the fifth step in detail. This way, we skip calculation-heavy parts of the derivation, and focus only on setting parameters so that our unitary transformation truly becomes the SWAP-gate.

Due to the coupling between the qubits and the magnetic field acting upon them, we arrive at the following Hamiltonian:

$$H(t) = \begin{pmatrix} -J_z + B_+ & 0 & 0 & -J_- \\ 0 & J_z + B_- & -J_+ & 0 \\ 0 & -J_+ & J_z - B_- & 0 \\ -J_- & 0 & 0 & J_z - B_+ \end{pmatrix}, \tag{6.12}$$

where $B_\pm(t) = B^1(t) \pm B^2(t)$ and $J_\pm(t) = J_x(t) \pm J_y(t)$. The quantities B^1 and B^2 represent the magnetic fields acting upon a 2-dimensional plane, whereas the quantities $J_{x,y,z}$ represent the coupling between the qubits along the respective axes. Furthermore, for the finite energy constraint, we assume that

$$B_+^2 + B_-^2 + J_x^2 + J_y^2 + 2J_z^2 = 2\omega^2, \tag{6.13}$$

where ω is some real number.

It can be shown that the quantities $B_\pm(t)$ and $J_\pm(t)$ should be given by

$$B_\pm(t) = B_{0\pm} \cos(2(\gamma_\pm t + \psi_\pm)) \tag{6.14}$$

and

$$J_\pm(t) = \mp B_{0\mp} \sin(2(\gamma_\mp t + \psi_\mp)) \tag{6.15}$$

to ensure time-optimal evolution. Here, $B_{0\pm}$, $\gamma_{0\pm}$ and $\psi_{0\pm}$ are to be determined parameters.

Since we now effectively have obtained $H(t)$, the related $U(t)$ can also be found. After some calculations,[18] we obtain that the optimal unitary transformation is

[18] See the source text [5] for full details.

given by

$$U(t) = \begin{pmatrix} e^{iJ_z t}(\alpha_{0+} + i\alpha_{z+}) & 0 & 0 & e^{iJ_z t}(\alpha_{y+} + i\alpha_{x+}) \\ 0 & e^{-iJ_z t}(\alpha_{0-} + i\alpha_{z-}) & e^{-iJ_z t}(\alpha_{y-} + i\alpha_{x-}) & 0 \\ 0 & e^{-iJ_z t}(-\alpha_{y-} + i\alpha_{x-}) & e^{-J_z t}(\alpha_{0-} - i\alpha_{z-}) & 0 \\ e^{iJ_z t}(-\alpha_{y+} + i\alpha_{x+}) & 0 & 0 & e^{iJ_z t}(\alpha_{0+} - i\alpha_{z+}) \end{pmatrix}, \tag{6.16}$$

where

$$\alpha_{0\pm}(t) = \cos(\gamma_\pm t)\cos(\Omega_\pm t) + \frac{\gamma_\pm}{\Omega_\pm}\sin(\gamma_\pm t)\sin(\Omega_\pm t) \tag{6.17a}$$

$$\alpha_{x\pm}(t) = \frac{\pm B_{0\pm}}{\Omega_\pm}\sin(\Omega_\pm t)\sin(\gamma_\pm t + 2\psi_\pm) \tag{6.17b}$$

$$\alpha_{y\pm}(t) = \pm\left(\sin(\gamma_\pm t)\cos(\Omega_\pm t) - \frac{\gamma_\pm}{\Omega_\pm}\cos(\gamma_\pm t)\sin(\Omega_\pm t)\right) \tag{6.17c}$$

$$\alpha_{z\pm}(t) = \frac{-B_{0\pm}}{\Omega_\pm}\sin(\Omega_\pm t)\cos(\gamma_\pm t + 2\psi_\pm), \tag{6.17d}$$

and $\Omega_\pm = \sqrt{B_{0\pm}^2 + \gamma_\pm^2}$. In the above equations, we have also introduced J_z and χ as constants which are to be determined. Together with the other parameters, they have to ensure that

$$U_{\text{SWAP}} = e^{-i\chi}U(T), \tag{6.18}$$

which precisely is step 5. of Plan 6.2. In total, we have eight constants for eight functions – this is doable.

The equation for $U(t)$, together with the expressions (6.17), is quite a horrendous set to attempt to solve. Fortunately, there is a way to pick the various constants such that the relation (6.18) can hold. This is the topic of the following proposition.

Proposition 6.3. *Regarding the constants so that $e^{-i\chi}U(T) = U_{SWAP}$, where $U(t)$ is as in equation* (6.16)*, we have that*

1. $B_{0+} = 0$
2. $\gamma_- = 0$
3. $B_{0-}T = \frac{\pi}{2}(1 + 2p)$
4. $2\psi_- = \frac{\pi}{2}(1 + 2q)$
5. $J_z T = \frac{-\pi}{4}[1 - 2(p + q) - 4(m - n)]$
6. $\chi = \frac{-\pi}{4}[1 - 2(p + q) + 4(m + n)]$

Here, m, n, p and q are arbitrary integers and the end time T is yet to be determined.

Proof. We shall only show that using the given values, the SWAP-gate is obtained – not actively derive the values ourselves. Though the latter would be instructive, it would also be time-consuming and distract from the overall use of the quantum brachistochrone.

Having said this, let it be seen that if we implement the values under points 1. through 4. into the functions (6.17) we obtain

$$\begin{array}{ll} \alpha_{0+} = 1 & \alpha_{0-} = 0 \\ \alpha_{x+} = 0 & \alpha_{x-} = (-1)^{p+q+1} \\ \alpha_{y+} = 0 & \alpha_{y-} = 0 \\ \alpha_{z+} = 0 & \alpha_{z-} = 0 \end{array}, \tag{6.19}$$

upon evaluating the functions at $t = T$. Thus, $U(T)$ becomes

$$U(T) = \begin{pmatrix} e^{iJ_zT} & 0 & 0 & 0 \\ 0 & 0 & ie^{-iJ_zT}(-1)^{p+q+1} & \\ 0 & ie^{-iJ_zT}(-1)^{p+q+1} & 0 & 0 \\ 0 & 0 & 0 & e^{iJ_zT} \end{pmatrix}. \tag{6.20}$$

Then, since we know the relation $e^{-i\chi}U(T) = U_{\text{SWAP}}$ holds by definition, we multiply the obtained expression for $U(T)$ by $e^{-i\chi}$ and then compare matrix elements so we can determine J_zT and χ.

Comparison of the matrix elements so reveals

$$\exp(i(J_zT - \chi)) = 1 \tag{6.21}$$

and

$$\exp(i(-J_zT - \chi + \pi(p+q) + \pi/2)) = 1. \tag{6.22}$$

Substituting the given values for J_zT and χ under 4. and 5, respectively, will yield that the equalities (6.21) and (6.22) hold. □

Notice that in Proposition 6.3, we sneakily introduced the end time T into our calculations so as to be able to say what happens at that time. Notice that as the physical interpretation of T is the optimal time in which the SWAP-gate will have been completed, it is only fitting to dub this end time T_{SWAP} instead. Together with the globally set constants – $B_{0+} = 0$, $\gamma_- = 0$ and $2\psi_- = \frac{\pi}{2}(1 + 2q)$ – we are able to distill an explicit expression for T_{SWAP}.

Proposition 6.4. *We have that the optimal time T_{SWAP} is given by*

$$\omega T_{\text{SWAP}} = \pi\sqrt{3}/4. \tag{6.23}$$

Proof. We utilize the values found in points 1, 2. and 4. of Proposition 6.3 together with the expressions in (6.14) and (6.15) to find

$$B_+ = B_- = J_- = 0 \text{ and } J_+ = -B_{0-}(-1)^q. \tag{6.24}$$

Filling these values into (6.13), we so obtain

$$B_{0-}^2 + 2J_z^2 = 2\omega^2 \Rightarrow \frac{1}{\pi^2}\left(8(B_{0-}T_{\text{SWAP}})^2 + 16(J_zT_{\text{SWAP}})^2\right) = \left(\frac{4\omega T_{\text{SWAP}}}{\pi}\right)^2. \tag{6.25}$$

We can now use the values found under points 3. and 5. of Proposition 6.3 by swapping[19] them into (6.25). The result is the lovely

$$\left(\frac{4\omega T_{\mathrm{SWAP}}}{\pi}\right)^2 = 2(1+2p)^2 + (1-2(p+q)-4(m-n))^2. \tag{6.26}$$

It follows that since we are looking for the shortest time duration, that we need to find m, n, p, q such that the right-hand side is minimized. This is the case for the choice of $p = 0$ and $q = -2(m-n)$, as can be readily verified.[20] With this choice, the right-hand side will equal 3, so that

$$\left(\frac{4\omega T_{\mathrm{SWAP}}}{\pi}\right)^2 = 3 \implies \omega T_{\mathrm{SWAP}} = \pi\sqrt{3}/4. \tag{6.27}$$

This proves the proposition. □

In addition, it can be shown that the Hamiltonian becomes

$$H = \frac{\omega(-1)^q}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{6.28}$$

upon finding $J_x = J_y = 2J_z = (-1)^{q+1}2\omega/\sqrt{3}$. Notable is that the Hamiltonian is constant, just as in the solution of the unconstrained quantum brachistochrone problem.

[19] Choice of word intentional.

[20] The choice for p minimizes the first term, whilst the subsequent choice for q then minimizes the second term.

7 Conclusion

We have come to the end of this thesis. All in all, quite a lot of material has been treated in the sixty-or-so pages which span this thesis.

In the first section, we discussed preliminaries from the calculus of variations – in particular the functional derivative – and the theory of Lagrange multipliers and its application for us.

Following this, we looked at the geometry of quantum states in order to derive an expression for the distance between them – the Fubini-Study metric. This was implemented together with various constraints in order to construct the quantum action, from which we could derive the equations of motion.

These were derived in the subsequent section. Furthermore, we managed to solve the equations of motion when dropping all but the "finite energy" constraint in closed form, yielding an optimal transition time of $\omega T = \cos^{-1} |\langle\psi_i|\psi_f\rangle|$; we also determined the associated state vector and Hamiltonian, which can be found in equations (4.92) and (4.93), respectively. We also solved the constrained version of the quantum brachistochrone problem, yielding several applicable formulae, particularly those the topic of Proposition 4.23.

This latter result was used to solve a particular example: a spin-1/2 particle under the influence of an x-y plane constrained magnetic field. We fully solved this example, i.e. we found the optimal transition time and the optimal state as a function of time, which results are in Proposition 5.10 and Theorem 5.9, respectively. Visualizations of the various solutions were also made, using the Bloch sphere.

Finally, we considered an application of the quantum brachistochrone to quantum computing. A brief introduction was given on the topic, after which an example was worked of the optimal unitary transformation to effect a given quantum gate.

As closing words, I would like to say that the research project and its resultant thesis were a success.

Topics for further research

In researching this problem, I have had to leave some avenues unexplored for the sake of time or brevity of my thesis. These should not be left unexplored, though, in my opinion.

For one, one could consider a different type of physics in which the quantum brachistochrone plays. That is, one could work with, for instance, the Klein-Gordon equation instead of the Schrödinger equation, and see in how so far the problem remains similar and/or solvable.

Another would be to investigate the link between quantum computing and the quantum brachistochrone more. For instance, to research whether companies/institutes working on quantum computers implement the theory learned from the quantum brachistochrone problem to optimize state transitions.

Additionally, one could investigate the relationship between the found results

for the (globally) optimal time of transitions and the uncertainity relation connecting energy uncertainty and time.

Acknowledgments

I thank all those who helped me write this thesis, whether little or lots. However, some special thanks are in order.

I would like thank Prof. Dr. Waalkens, my mathematics supervisor, for the frequent meetings we had discussing this thesis. These near-weekly meetings gave me a better insight into the underlying physics and mathematics of the problem, which explanations often helped me through tough spots in the research.

I would also like to thank C. Hermans, for writing his thesis on the same topic. With the help of his thesis, the necessary calculations for solving the quantum brachistochrone were much more manageable than they otherwise would have been. As a result, I was able to research more broadly into the application and the derivation of the quantum action.

Furthermore, I would like to thank dr. Hosoya and dr. Okudaira, two of the authors of [4], for taking the time to respond to some of my questions I had about their paper. Their answers gave me a better insight into the contents of their paper and how to find additional literature.

References

[1] Anandan, J., and Aharanov, Y. Geometry of Quantum Evolution. *Physical Review Letters*, 65 (October 1990). 1697.

[2] Bengtsson, I., and Zyczkowski, K. Geometry of Quantum States, March 2015. Chapters 3, 4 and 5.

[3] Brody, D., and Hook, D. On Optimum Hamiltonians for State Transformations. *Journal of Physics A: Mathematical and General*, 39 (March 2006). L167; corrected in J. Phys. A: Math. Theor. (40) 2007 10949.

[4] Carlini, A., Hosoya, A., Koike, T., and Okudaira, Y. Time-Optimal Quantum Evolution. *Physical Review Letters*, 96 (February 2006). 060503.

[5] Carlini, A., Hosoya, A., Koike, T., and Okudaira, Y. Time-optimal Unitary Transformations. *Physical Review A*, 75 (April 2007). 042308.

[6] Dym, C. L., and Shames, I. H. *Solid Mechanics: A Variational Approach*. Springer, 2013. Chapter 2.

[7] Foot, C. J. *Atomic Physics*. Oxford University Press, 2014. Chapter 7.

[8] Hermans, C. The Brachistochrone Problem: From Euler to Quantum. Bachelor's thesis, Delft University of Technology, July 2017.

[9] Rieffel, E., and Polak, W. *Quantum Computing: A Gentle Introduction*. MIT Press, 2011. Section I.

[10] van der Schaft, A. *Calculus of Variations and Optimal Control*. University of Groningen, 2017. Self-published textbook/syllabus.