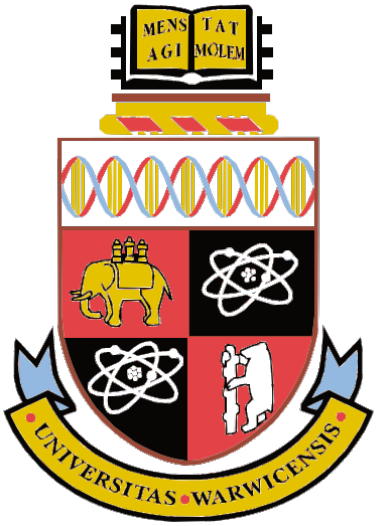

Adiabatic theory of a two-level quantum system

by

Gabriel Rémond-Tiedrez

MA4K9 Dissertation

Submitted to The University of Warwick

Mathematics Institute

March, 2023

Contents

1 Introduction

1.1 A qubit of history and motivation

Information theory dates back to the 1920's, and is the basis on which computers were theorised as useful, then built, used and improved. Contemporary computers are orders of magnitude more powerful than the first, and yet they still have very real limitations. The approach taken by any classical – non-quantum – computer is discrete: it works with deterministic bits in 2 possible states of 0 or 1. As such, computers may only approximate continuous behaviours, which can be done within astounding accuracy but they remain a quantifiable error away. If more processing power is added, more accuracy can follow. However, we then come across a physical barrier: to increase processing power while keeping size of computers somewhat bounded, we require transistors to shrink, and if the current trend continues we will reach atomic or subatomic scales. Once this happens, we are in the realm of quantum mechanics, where the deterministic structure of the computer collapses.

In contrast, quantum computers consist not of bits, but of quantum bits, or qubits. One qubit is a complex linear combination of the 2 possible states for a classical bit, a 0 or a 1. Each coordinate component is a probability amplitude, where taking the modulus squared gives the probability of finding the qubit in the given state. Similarly, a system of n-qubits is a complex linear combination of all 2^n possible states for n classical bits. Where classical computers use discrete algorithms, quantum computers could simulate continuous models exactly. For more reading on quantum computing and its physical realisations, the reader is referred to [MM04, RX11][1].

Another application of quantum computing is for cryptography. For one, Shor's algorithm – see [Sho94] – can be used to factorise integers in polynomial time, an unachievable feat for classical computers. A working quantum computer given an adequate amount of qubits could therefore dismantle public-key cryptography. On the other hand, one could use quantum key distribution to ensure a sent key is secure. This can use the properties of superposition or entanglement, but either relies on the fact that measuring a quantum system induces detectable changes. Eavesdropping on the key exchange would then notify the communicators, who would abort the procedure. This is a very active area of research, so the reader may find many relevant articles, one of which is [BB14].

Before the main body of this report, I would like to make a couple of acknowledgements. I want to thank to my supervisor Prof. Vassili Gelfreykh for introducing me to this fascinating area of mathematics, as well as for all the help and all the time taken for this project. I also want to thank my friends and Kaijae for keeping me sane and everything else too.

[1] The pun "qubit" of history is also from [MM04], and was too good not to include.

1.2 Mathematical background

The usual environment of quantum mechanics is infinite dimensional spaces, but as the interest in quantum computing grows, finite dimensional quantum mechanics proves itself quite relevant. We now present useful notions and intuition from this theory: the following will be a selection of pertinent material in [Gel18], mostly from chapter 3. Throughout this dissertation, we consider the 2-dimensional Hilbert space $\mathbb{C}^2$, which in the context of quantum computing, should be thought of as the space of complex linear combinations of the classical bits 0 and 1: we have a single qubit.

Like in any physical system, there are 2 main objects of study: possible states that the system can be in, and observables on this system. In classical mechanics, states are described by $2n$-tuples $(q_1, \dots, q_n, p_1, \dots, p_n) \in \mathbb{R}^{2n}$, and observables are smooth functions $f : \mathbb{R}^{2n} \to \mathbb{R}$. A standard example is the total energy of the system, which could be given by the sum of potential and kinetic energies. In quantum mechanics, states are described by density matrices and observables are general Hermitian matrices.

1.2.1 Density matrices

A density matrix describes a state for the system, and is a Hermitian matrix $M \in \mathbb{C}^{2\times 2}$, such that

1. $\langle M\psi, \psi\rangle \geq 0$ for all $\psi \in \mathbb{C}^2$,
2. $\operatorname{Tr} M = 1$.

Note that all 2×2 Hermitian matrices can be written as $M = \lambda_1\pi_1 + \lambda_2\pi_2$, where the λ_i's are the eigenvalues of M and the π_i's are orthogonal projections onto the corresponding eigenvector. A projection is defined as a Hermitian operator π such that $\operatorname{Tr}\pi = 1$ and $\pi^2 = \pi$. For density matrices, condition 1 on M implies that the eigenvalues $\lambda_{1,2} \geq 0$, and condition 2 gives $\lambda_1 + \lambda_2 = 1$.

Orthogonal projections are very convenient objects to work with, so we present some useful properties. Any projection onto a 1-dimensional subspace can be written as $\pi_\psi(\varphi) = \langle\varphi, \psi\rangle\psi$, for[2] some unit vector $\psi \in \mathbb{C}^2$, which can be written more compactly using Dirac notation as $\pi_\psi = |\psi\rangle\langle\psi|$. However, this choice of unit vector is not unique, since we can multiply by any phase factor $e^{i\theta}$ without changing the projection[3], so the relation between ψ and π_ψ is not 1-to-1, but rather $\mathbb{S}^1$-to-1. Observe that ψ is an eigenvector of π_ψ corresponding to eigenvalue 1, and its orthogonal complement $\psi^\perp$ an eigenvector with eigenvalue 0. We can obtain $\psi^\perp$ using the explicit formula available in 2 dimensions, giving $\psi^\perp = \begin{pmatrix} -\overline{\psi}_2 \\ \overline{\psi}_1 \end{pmatrix}$.

[2] We define $\langle\varphi, \psi\rangle = \varphi_1\overline{\psi}_1 + \varphi_2\overline{\psi}_2$, that is the inner product is conjugate linear in the second component.
[3] The factor must have modulus 1, otherwise the vector would no longer have norm 1.

We have 2 possibilities for density matrices: either $\lambda_1 = 1$ and $\lambda_2 = 0$, or both are non-zero. We call the state pure in the former case, and mixed in the latter. If M is pure, it is a projection, so we write $M = \pi_\psi$ and we refer to ψ as the wave function associated with the quantum state. If M is mixed, we can write $M = \lambda_1\pi_\psi + \lambda_2\pi_{\psi^\perp}$, noting that M has orthogonal eigenvectors as it is Hermitian. The quantum state being mixed is interpreted as having probabilistic data: the system is described by the wave function ψ with probability λ_1, or by $\psi^\perp$ with probability $\lambda_2 = 1 - \lambda_1$. This interpretation says that the pure state π_ψ can be described by the wave function ψ with probability 1.

1.2.2 Observables and their interactions

Observables are any Hermitian matrix $A \in \mathbb{C}^{2\times 2}$. We can make observables interact with each other, by use of a quantum Poisson bracket, but for the purpose of this dissertation, we will be looking at their interaction with quantum states. Suppose the system is in the state described by density matrix M and consider the observable A. We can then calculate the expectation of A, which we define by[4]

$$\mathbb{E}_M[A] = \mathrm{Tr}(MA).$$

To justify the choice of terminology and notation, let us first write $A = \lambda_1\pi_1 + \lambda_2\pi_2$, where the λ_i's are eigenvalues of A, and π_i orthogonal projections onto the corresponding eigenspaces, as before. We now consider the random variable a, taking value λ_i with probability $\mathrm{Tr}(M\pi_i)$. One can check that the random variable a has expectation $\mathbb{E}_M[A]$.

An interesting observation is that density matrices can be considered as observables. For the rest of this dissertation, we will consider projections instead of general density matrices. Let π_ψ, π_φ be projections for some unit vectors $\psi, \varphi \in \mathbb{C}^2$. Suppose that the system is in the quantum state described by π_ψ and consider π_φ as an observable. We calculate the expectation

$$\mathbb{E}_{\pi_\psi}[\pi_\varphi] = \mathrm{Tr}(\pi_\psi\pi_\varphi) = \mathrm{Tr}(|\psi\rangle\langle\psi||\varphi\rangle\langle\varphi|),$$

which yields $\mathbb{E}_{\pi_\psi}[\pi_\varphi] = |\langle\psi, \varphi\rangle|^2$. If we consider the random variable a defined above, it takes value 1 with probability $\mathrm{Tr}(\pi_\psi\pi_\varphi) = |\langle\psi, \varphi\rangle|^2$, and takes value 0 with probability $1 - |\langle\psi, \varphi\rangle|^2$. This should be interpreted in the following way: which wave function better describes the system between φ or $\varphi^\perp$ given that the system is in state π_ψ.

Observe that multiplying ψ or φ by any phase factor does not change the probability, another showcase of the relation between the wave function and the quantum state. At first glance, the overall phase of a wave function seems to have no physical meaning. However, if one considers the superposition of 2 wave functions with a non-zero relative phase, a phenomenon called quantum interference occurs, which is physically measurable[5].

[4] As A is Hermitian, one has $\mathbb{E}_M[A] = \mathrm{Tr}(MA^*)$, which gives the usual inner product on $\mathbb{C}^{2\times 2} \cong \mathbb{C}^4$.

[5] The reader is referred to section 2.2.1 of [CJ04] for a short discussion on this.

1.2.3 Time evolution

With the quantum states and observables in place, one can now study the time evolution of quantum systems. Just as in classical mechanics, one considers an observable H called the Hamiltonian, which is interpreted as describing the energy of the system. Note that for our 2-level quantum system, this is a 2×2 Hermitian matrix. As we previously saw, there are 2 almost equivalent ways to describe pure quantum states: orthogonal projections onto 1-dimensional subspaces, and unit vectors, referred to as wave functions. We can therefore describe the time evolution of quantum systems by looking at the behaviour of either. One has the two linear equations:

$$i\hbar\dot{\psi} = H\psi, \tag{1}$$

for unit vectors $\psi \in \mathbb{C}^2$, and

$$i\hbar\dot{\pi} = [H, \pi], \tag{2}$$

for projections π, where $[A, B] = AB - BA$ denotes the commutator of 2 matrices, and $\hbar$ is Planck's constant. The former is the well-known Schrödinger equation and the latter the von Neumann equation, both of which are non-autonomous as H is a function of time[6]. Another equation of interest is the Schrödinger equation for unitary matrices:

$$i\hbar\dot{K} = HK, \tag{3}$$

for $K \in SU(2)$. Let ψ be a solution to (1). If H is traceless, $\psi^{\perp}$ also solves (1), and considering the matrix with columns $\psi, \psi^{\perp}$, that is $K = \begin{pmatrix} \psi_1 & -\overline{\psi}_2 \\ \psi_2 & \overline{\psi}_1 \end{pmatrix}$, K is unitary, has determinant 1 from $\|\psi\| = 1$, and solves (3). Solving the equation for unit vectors and special unitary matrices is therefore completely equivalent in this low dimension.

If one adds the condition $K(s, s) = I_2$, one then writes K as a matrix exponential $K(t, s) = \exp\left(-i/\hbar \int_s^t H\right)$, which is called the quantum evolution operator. We can then write any solution ψ to (1) as $K(t, 0)\psi_0$ for some constant unit vector ψ_0. Note that in any dimension larger than 2, more work is needed to obtain the evolution operator, but we only require a solution of the Schrödinger equation with $\psi(s, s) = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$. Throughout this dissertation, we will be exploiting the low dimension of the problem to obtain different viewpoints, using its simplicity and not looking for a generalisable approach, unlike in [BT05a].

1.3 Quantum adiabatic theory

Most quantum systems deal with a continuum of possible states, requiring infinite dimensions, functional analytic approaches and spectral theory for linear operators. This is the context for the usual adiabatic quantum theory.

[6] Note that $\hbar$ is not strictly necessary for mathematical purposes as a quick change of variables normalises it to 1.

Adiabatic processes naturally occur in many areas of physics, where a system splits into different time-scales. When one has 2 time-scales, one is called microscopic and the other macroscopic, the former being "fast" and the latter "slow". One encounters such systems in thermodynamics, as well as celestial, nuclear or quantum mechanics.[7].

The quantum adiabatic theorem was first proved by Born and Fock in [BF28], then refined by Kato in [Kat50], and an extensive explanation and discussion can be found in [Teu03]. Another reading suggestion is [HJ05], which goes over the history and different approaches to the problem. Simply put, for a slowly changing Hamiltonian, the spectral subspaces of the Hamiltonian are approximately invariant under the time evolution. In the finite dimensional case, spectral subspaces are instantaneous eigenspaces. We refer to the case where a solution goes from one such subspace to another as a transition. In a general infinite dimensional setting, one requires a spectral gap condition on the Hamiltonian H, which in finite dimensions is simplified to separating eigenvalues by some gap, uniformly for all times.

For the specific case of 2-dimensional systems, Berry heuristically calculates the probability of transitions between eigenstates of H, and presents the idea that it does not on the precise time-dependence structure of the Hamiltonian, see [Ber90] sections 1 and 4. Hagedorn and Joye have rigorously shown the statement for a restricted class of Hamiltonians in [HJ04], and Betz and Teufel for a slightly more general class in [BT05a]. Betz and Teufel's argument revolves around diagonalising H to get exponentially small coupling terms, i.e. off-diagonal elements of H. They then solve the Schrödinger equation for this new Hamiltonian and compute the transition amplitude.

The natural starting point for analysing the relevant Schrödinger equation is to express it in a basis of eigenvectors of H, called the adiabatic basis. However, not only does this not exactly diagonalise H, but this basis is also suboptimal for studying transitions, as one only gets a rather primitive bound in terms of the adiabatic parameter. To remedy this, one looks for a different, so-called superadiabatic basis. There are a few ways of obtaining such a basis, one is by finding eigenvectors of H, expressing H in this basis, finding eigenvectors of the new Hamiltonian and iterating the process a number of times. Another method uses the low dimension of the problem: if a solution to the equation is given, it can be used to diagonalise H, which makes finding solutions as easy as integrating. One therefore finds an approximate solution, written as a power series in terms of the adiabatic parameter, and uses it to give an almost diagonal H. This method is what we use in this dissertation.

The iterative process, or the series, will in general not converge, which is why we get a possibility of transition. This encourages another essential aspect of these bases, which is choosing an optimal one for every given value of the adiabatic parameter. This crucial point is discussed in [Ber90] section 3, and [BT05a] sections 2, 3 and 5.

[7] An example of the last is a particle in an electric field that one slowly turns on.

In the following sections we discuss the work of Betz and Teufel [BT05a]. We follow the paper very closely, the main differences being a different order of presentation, the explicit mention of the von Neumann equation and the introduction of the Pauli spin matrices before obtaining the recursion.

2 Setting up the problem

We analyse solutions to the 2-level quantum system

$$i\varepsilon\dot{\psi} = H\psi, \tag{4}$$

in the adiabatic limit $\varepsilon \to 0$, with dots denoting t-derivatives throughout. Here $H(t)$ is a Hermitian operator on $\mathbb{C}^2$ depending on $t \in \mathbb{R}$, and $\psi : \mathbb{R} \to \mathbb{C}^2$ is a wave function, that is $\|\psi(t)\| = 1$ for all t. One can obtain (4) by considering the usual Schrödinger equation with a slow change in the Hamiltonian:

$$i\hbar\psi'(\tau) = H(\delta\tau)\psi(\tau)$$

for $\delta > 0$ small, with prime denoting a τ-derivative. Then the change of variables $t = \delta\tau$, $\varepsilon = \hbar\delta$, yields exactly (4). This changes the time scale from microscopic to macroscopic, while normalising the constant $\hbar$ to 1, and is what [Ber90] starts with. We can also write the von Neumann equation (2) with this adiabatic parameter to get

$$i\varepsilon\dot{\pi} = [H, \pi]. \tag{5}$$

2.1 Simplifying the Hamiltonian

We consider how H changes under special unitary change of variables. More precisely, if $\tilde{\psi} = U^*\psi$ for some $U \in SU(2)$ and ψ solves (4), then $\tilde{\psi}$ solves

$$i\varepsilon\dot{\tilde{\psi}} = \tilde{H}\tilde{\psi},$$

where the new Hamiltonian is

$$\tilde{H} = U^*HU - i\varepsilon U^*\dot{U}, \tag{6}$$

which stays traceless precisely if $\det U = 1$. Observe that this Hamiltonian is zero if U solves the Schrödinger equation for matrices, which could be problematic, cf section 3.3. If we let $\tilde{\pi} = U^*HU$, then (5) sees the same change in H, that is $\tilde{\pi}$ solves

$$i\varepsilon\dot{\tilde{\pi}} = [\tilde{H}, \tilde{\pi}].$$

We now simplify H without losing generality. Note that any 2×2 Hermitian matrix can be written as

$$H = \begin{pmatrix} s+z & x\exp(-i\phi) \\ x\exp(i\phi) & s-z \end{pmatrix},$$

where s, x, z, ϕ are all functions of time t. Now observe that subtracting any multiple of I_2 to H will not change the von Neumann equation, as I_2 commutes trivially. We can therefore assume that $s \equiv 0$[8]. Now consider the special unitary matrix

$$U = \begin{pmatrix} \exp(\frac{-i}{2}\phi) & 0 \\ 0 & \exp(\frac{i}{2}\phi) \end{pmatrix},$$

which gives

$$\tilde{H} = \begin{pmatrix} \tilde{z} & \tilde{x} \\ \tilde{x} & -\tilde{z} \end{pmatrix}$$

by (6), where $\tilde{x} = x$ and $\tilde{z} = z - \frac{1}{2}\varepsilon\dot{\phi}$.

We can then rewrite this through a more geometric argument. One can think of $\tilde{x}, \tilde{z}$ as Cartesian coordinates in "Hamiltonian space", as Berry puts it. Using standard polar coordinates to write $\tilde{x} = \rho\sin\theta, \tilde{z} = \rho\cos\theta$ for some functions ρ, θ of time, we have

$$\tilde{H} = \rho \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

We now assume $\rho \equiv 1/2$, which can be done by changing to the natural time scale

$$\tau(t) = 2\int_0^t \rho(u)\mathrm{d}u.$$

The Hamiltonian is therefore simplified to

$$H = \frac{1}{2}\begin{pmatrix} \cos\theta(t) & \sin\theta(t) \\ \sin\theta(t) & -\cos\theta(t) \end{pmatrix}, \tag{7}$$

for some function θ of time, called the coupling.

One can consider a choice of unit eigenvectors of H, called the adiabatic basis, and consider the corresponding special unitary matrix,

$$U_0 = \begin{pmatrix} \cos\theta/2 & \sin\theta/2 \\ \sin\theta/2 & -\cos\theta/2 \end{pmatrix}.$$

[8]This argument does not quite work for the Schrödinger equation, but one can apply a special unitary transformation to H to ensure $s \equiv 0$, see [Ber90], appendix A.

One uses (6) to obtain that the Hamiltonian in the adiabatic basis is

$$H_a = U_0^* H U_0 - i\varepsilon U_0^* \dot{U}_0 = \begin{pmatrix} \frac{1}{2} & \frac{i\varepsilon}{2}\dot{\theta} \\ -\frac{i\varepsilon}{2}\dot{\theta} & -\frac{1}{2} \end{pmatrix}. \tag{8}$$

Throughout this dissertation, we will assume that the coupling has

$$\dot{\theta} = \frac{2\gamma t_c}{t^2 + t_c^2}, \text{ that is } \theta = 2\gamma \arctan(t/t_c) \tag{9}$$

for $\gamma, t_c > 0$, just as in [BT05a][9]. This is the only restriction on the Hamiltonian, but one can show that the coupling is close to this in a generic class of Hamiltonians, see [BT05b]. From now onwards, we will consider the Hamiltonian as in (8). This is the first notable difference from [BT05a], as Betz and Teufel work in the standard basis and keep the change of basis matrix U_0 in all their calculations.

2.2 Scattering

We can split the Hamiltonian into 2 parts: $H = H_0 + V_\varepsilon$ where the former is independent of ε and the latter depends of ε. Explicitly, we have $H_0 = 1/2 \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $V_\varepsilon = \varepsilon\dot{\theta}/2 \left(\begin{smallmatrix} 0 & i \\ -i & 0 \end{smallmatrix}\right)$. This is a common start in perturbation theory, and we can view (4) as a perturbation of the autonomous equation $i\varepsilon\dot{\psi} = H_0\psi$, which can easily be solved explicitly as H_0 is constant and diagonal. Moreover, the perturbation V_ε decays at infinity, since $\dot{\theta} \to 0$ as $|t| \to \infty$, so we have additional structure to work with.

Let ψ be a solution to (4), the perturbed problem. For t large, V_ε is small and ψ behaves like a solution to the unperturbed case. More precisely, we can find some solutions $\psi_\pm$ to the unperturbed problem, such that $\psi \approx \psi_\pm$ as $t \to \pm\infty$, where one thinks of ψ_- as the behaviour of ψ for t large and negative, and ψ_+ as the behaviour of ψ for t large and positive. We can then define a map S called the scattering map, that takes ψ_- to ψ_+, and which will be computed near the end of this report.

The physical relevance of the scattering is clearer when considering eigenstates of the Hamiltonian. Let v_1, v_2 be the eigenvectors of H_0, with $v_1 = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right), v_2 = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ and say that v_2 represents the ground state and v_1 the excited state. Suppose that we start at the ground state for large negative times and we want to compute the probability of transition to the excited state at large positive times. In our notation above, we have $\psi_- = v_2$ and the probability of transition will be $|\langle \psi_+, v_1 \rangle|^2$, where $\langle \psi_+, v_1 \rangle$ is called the transition amplitude. The main goal of this report is to compute this transition amplitude, which we will obtain by computing the scattering map.

We now explain the structure and main arguments. We first show that having one solution allows diagonalising the Hamiltonian, and similarly an approximate solution gives an

[9] Betz and Teufel write $\dot{\theta}$ without the factor of 2, but in all their calculations they use $\dot{\theta}$ as written here.

almost diagonal Hamiltonian. In order to obtain this approximate solution, we introduce an equivalent formulation of the von Neumann equation on the unit 2-sphere. We then use asymptotic series, analyse and solve the scalar recursion obtained through them; this is the most time consuming part. Next, we go back to projections and obtain a so-called superadiabatic unitary change of basis, in which the Hamiltonian has off-diagonal elements of order ε^{n+1}. Then choosing n optimally as a function of ε yields that the off-diagonal elements are in fact exponentially small, which is used to find the evolution operator for the Schrödinger equation, in turn giving the scattering map and relevant transition amplitude.

3 Overture

3.1 One for all

This problem has the crucial property that finding one solution enables finding all of them. More precisely, let π be a solution to (5) and let U be the special unitary matrix that diagonalises it to $P = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, that is $UP = \pi U$. Then changing the basis using U changes the Hamiltonian to be diagonal, which we express as $P\tilde{H}(1-P) = (1-P)\tilde{H}P = 0$. Note that checking that the Hamiltonian is diagonal generally requires more algebra than this, which only works for 2×2 matrices. We only compute

$$\begin{aligned} PU^*(HU - i\varepsilon\dot{U})(1-P) &= U^*\left(\pi HU - i\varepsilon\pi\dot{U} - \pi HUP + i\varepsilon\pi\dot{U}P\right), \\ &= U^*\left(\pi HU - \pi H\pi U + \pi[H,\pi]U\right), \\ &= 0, \end{aligned}$$

where we used $i\varepsilon\dot{U}P = i\varepsilon\pi\dot{U} + [H,\pi]U$ and $\pi^2 = \pi$. Solving the von Neumann equation is then as easy as integrating, and we can use U, U^* to send us back to the original basis. This property also means that approximate solutions will render the Hamiltonian "almost diagonal", intuition which will be made rigorous and used later.

3.2 On the unit 2-sphere

We have already seen 2 relevant equations when considering time evolution of quantum systems: the Schrödinger and von Neumann equations. In this 2-dimensional case, we can obtain 2 others equivalent to the von Neumann equation, with the added advantage of being easier to visualise. We explore one here and the other is discussed in appendix A.

Recall the von Neumann equation (5) for projections. Consider the Pauli spin matrices, defined by

$$\sigma_0 = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{10}$$

These form a basis for the vector space of Hermitian 2×2 matrices, and one useful property is their nice behaviour under commutation. Indeed, we have $[\sigma_x, \sigma_y] = 2i\sigma_z$, and applying 3-cycles to the symbols x, y, z preserves the equality while transpositions flip the sign. This basis therefore behaves very well under the relevant operation in (5).

One can write $\pi = \frac{1}{2}(I_2 + x\sigma_x + y\sigma_y + z\sigma_z)$ for some real numbers x, y, z, where the condition $\operatorname{Tr}\pi = 1$ is immediate and the condition $\pi^2 = \pi$ gives $x^2 + y^2 + z^2 = 1$, which is also equivalent to $\det\pi = 0$[10]. One also has $H = \frac{1}{2}(a\sigma_x + b\sigma_y + c\sigma_z)$ for some real numbers a, b, c, noting that $a = 0, b = -\varepsilon\dot{\theta}$ and $c = 1$ with (8). We can then consider the vectors $\mathbf{a} = (a, b, c)$ and $\mathbf{x} = (x, y, z)$, and check that the von Neumann equation transforms to

$$\varepsilon\dot{\mathbf{x}} = \mathbf{a} \times \mathbf{x}, \tag{11}$$

where $\times$ is the usual cross product in $\mathbb{R}^3$. Observe that $\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{x}\|^2 = 0$, so $\|\mathbf{x}\| = 1$ for all t, that is $\mathbf{x}$ stays on the unit sphere. This is commonly called the Bloch sphere, and is useful to visualise 2-level quantum systems. Higher dimensional analogues exist, but this process only really helps visualising in this simple case.

If we let A be the 3×3 real matrix such that $\mathbf{a}\times\mathbf{x} = A\mathbf{x}$, noting that A is anti-symmetric and traceless, we can define the matrix exponential $R = \exp(\frac{1}{\varepsilon}\int_s^t A)$. Then $RR^T = I_3$ and $\det R = 1$, so $R \in SO(3)$, the group of rotations of the 2-sphere. This is the evolution operator of the system, where any solution to (11) can be written as $R\tilde{\mathbf{x}}$ for some constant unit vector $\tilde{\mathbf{x}}$. This makes it clear that the behaviour of solutions is described by rotations, and one can observe that the axis of rotation will be given by the direction of $\mathbf{a}$ and the speed will depend on its magnitude. The picture one should have in mind for (11) is that of a sphere where solutions rotate rapidly around a slowly changing axis.

For a solution to the Schrödinger equation which starts at the ground state $v_2 = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$, the corresponding projection is $\pi = \left(\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right)$, which corresponds the South pole $(0, 0, -1)$ on the sphere, after using the Pauli spin matrices. A solution to (4) starting at v_2 at time $-\infty$ therefore corresponds to a solution to (11) starting at $(0, 0, -1)$. This is exactly the solution we will approximate in section 4.

One can use stereographic projection to send this onto the complex plane to obtain another equivalent formulation, as discussed in appendix A.

3.3 Asymptotic series

The equations of interest all require some perturbation theory, so we present an overview of the motivating ideas.

[10] Assuming $\operatorname{Tr}\pi = 1$, one can replace the condition $\pi^2 = \pi$ by $\det\pi = 0$ for 2 by 2 Hermitian matrices to be projections, but not for any larger matrix.

Consider (11), which we can rewrite as

$$\varepsilon\dot{\mathbf{x}} = (e_3 - \varepsilon\dot{\theta}e_2) \times \mathbf{x},$$

where e_1, e_2, e_3 is the canonical basis of $\mathbb{R}^3$. Writing the Hamiltonian in the adiabatic basis from the start allows us to explicitly split $\mathbf{a}$ into a term independent of ε and the other depending on ε. The standard idea in perturbation theory is to start with the easier part, that is the part independent of ε. Let $\mathbf{x}_0$ be such that $0 = e_3 \times \mathbf{x}_0$, we can then calculate how close $\mathbf{x}_0$ is to solving the equation, that is compute

$$\varepsilon\dot{\mathbf{x}}_0 - (e_3 - \varepsilon\dot{\theta}e_2) \times \mathbf{x}_0 = \varepsilon\dot{\theta}e_2 \times \mathbf{x}_0,$$

which is of order ε. Now consider $\mathbf{x}_1$ such that $\dot{\mathbf{x}}_0 = v \times \mathbf{x}_1 + \varepsilon\dot{\theta}e_2 \times \mathbf{x}_0$, and add a linear order correction to get $\mathbf{x}^{(1)} = \mathbf{x}_0 + \varepsilon\mathbf{x}_1$, observing that

$$\varepsilon\dot{\mathbf{x}}^{(1)} - (e_3 - \varepsilon\dot{\theta}e_2) \times \mathbf{x}^{(1)} = \varepsilon^2(\dot{\mathbf{x}}_1 + \dot{\theta}e_2 \times \mathbf{x}_1),$$

that is the error is now of order ε^2. One can keep iterating this process to continue increasing the order of the error, and consider a formal series of the form

$$\mathbf{x} = \sum_{k\geq 0} \varepsilon^k \mathbf{x}_k. \tag{12}$$

We will see in the following section that this sum diverges, so instead of considering the infinite object, we will consider finite sums, obtained after truncation. These divergent series are called asymptotic series, and we will define one for each equation previously discussed[11]. One of the crucial aspects of asymptotic series is the notion of optimal truncation, that is to truncate the sum at the smallest term, which will give the most information about the function of interest. We will discuss this in more detail in section 5.

Let us start with the sphere equation, and the other equations will receive a similar treatment. We consider solutions to (11) which we can write as power series in ε as in (12), where each $\mathbf{x}_k$ is independent of ε and the series commutes with differentiation. Substituting this into the equation and identifying coefficients of powers of ε yields:

$$0 = e_3 \times \mathbf{x}_0, \tag{13}$$

$$\dot{\mathbf{x}}_k = e_3 \times \mathbf{x}_{k+1} - \dot{\theta}e_2 \times \mathbf{x}_k, \text{ for } k \geq 0. \tag{14}$$

[11] At first glance, asymptotic series being divergent seems to make them impossible or useless to analyse, but they offer a beautifully rich theory, as presented in [Din73].

Note that we also have the condition $\|\mathbf{x}\| = 1$, meaning that we require

$$\|\mathbf{x}_0\| = 1, \tag{15}$$

$$\sum_{j=0}^{k+1} \langle \mathbf{x}_j, \mathbf{x}_{k+1-j} \rangle = 0, \text{ for } k \geq 0. \tag{16}$$

These imply that $\mathbf{x}_0 = \pm e_3$, and every subsequent $\mathbf{x}_k$ is obtained by solving the recursion.

The following will not be used, but are presented for the sake of completeness, and to give a more extensive picture of the theory. We look at the equivalent equation for a projection π, where the relevant equation is (5) with $H = \frac{1}{2}(\sigma_z - \varepsilon\dot{\theta}\sigma_y)$, where σ_y, σ_z are Pauli spin matrices as defined in (10). For solutions of the form

$$\pi = \sum_{k\geq 0} \varepsilon^k \pi_k,$$

we have

$$\begin{aligned} 0 &= [\sigma_z, \pi_0], \\ i\dot{\pi}_k &= \frac{1}{2}[\sigma_z, \pi_{k+1}] - \frac{1}{2}[\sigma_y, \pi_k], \text{ for } k \geq 0. \end{aligned}$$

The conditions required here are $\pi^2 = \pi$ and $\operatorname{Tr}\pi = 1$, which yield

$$\begin{aligned} \pi_0^2 &= \pi_0, && \text{and } \operatorname{Tr}\pi_0 = 1, \\ \pi_{k+1} &= \sum_{j=0}^{k+1} \pi_j \pi_{k+1-j}, && \text{and } \operatorname{Tr}\pi_{k+1} = 0 \text{ for } k \geq 0. \end{aligned}$$

Note that π_0 only has 2 possibilities, either $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. While the recursion is equivalent to (14), the conditions change since $\operatorname{Tr} = 1$ is immediate when considering $\mathbf{x}$, but require checking for π. Note that these are completely equivalent to (21) and (22) from [BT05a], with the only difference being that we wrote the Hamiltonian in the adiabatic basis to start with.

Lastly, we return to the Schrödinger equation (4) with the condition $\|\psi\| = 1$, which will require a slight change of approach. Indeed, trying

$$\psi = \sum_{k\geq 0} \varepsilon^k \psi_k$$

yields $0 = \frac{1}{2}\sigma_z\psi_0$ with $\|\psi_0\| = 1$, a contradiction since σ_z has full rank. We therefore try the ansatz

$$\psi = e^{-\alpha i/\varepsilon} \sum_{k\geq 0} \varepsilon^k \psi_k$$

for some function of time α, which gives

$$\dot{\alpha}\psi_0 = \frac{1}{2}\sigma_z\psi_0,$$
$$\dot{\alpha}\psi_{k+1} + i\psi_k = \frac{1}{2}\sigma_z\psi_{k+1} - \frac{1}{2}\dot{\theta}\sigma_y\psi_k, \text{ for } k \geq 0.$$

This difference might seem strange, but recall that the property that makes the Schrödinger and von Neumann equations not equivalent is the phase of ψ, which projections ignore. It is therefore a good sanity check that this is also the only difference in the corresponding series.

As mentioned earlier, these asymptotic series will be more useful when truncating them. To this end, consider $n \in \mathbb{N}$ and let $k = 0, 1, \ldots, n$. We construct the truncated sum, say for the sphere equation, by

$$\mathbf{x}^{(n)} = \sum_{k=0}^{n} \varepsilon^k \mathbf{x}_k.$$

The notion of optimal truncation is to fix ε and find n_ε for which the last term $\varepsilon^{n_\varepsilon}\mathbf{x}_{n_\varepsilon}$ is minimal, then taking the corresponding sum $\mathbf{x}^{(n_\varepsilon)}$.

We conclude this section with 2 observations. First, note that each series gives 2 possible starting points, each corresponding to an eigenspace of H_0 – the part of the Hamiltonian independent of ε. This is most obvious with the Schrödinger equation, where ψ_0 is an eigenvalue of $H_0 = (1/2)\sigma_z$ with eigenvalue $\dot{\alpha}$. For the von Neumann equation on projections, one can notice that each π_0 is a projection onto an eigenspace of $(1/2)\sigma_z$.

Finally, observe that only examining solutions of this form is quite a considerable restriction. Indeed, general solutions to the von Neumann equation (5) contain a rapidly oscillating term $e^{-i/\varepsilon}$, which is not expandable in positive powers of ε. Not allowing generak solutions has the consequence that we do not allow all possible $SU(2)$ matrices to change the Hamiltonian as in (6). Under a change of basis by $U \in SU(2)$, the new projection is $\tilde{\pi} = U^*\pi U$, so if we want $\tilde{\pi}$ to be a power series in ε, we require U to be expandable in powers of ε as well. This means that U cannot solve the Schrödinger equation. For a contradiction, suppose it does. Then the first column of U solves the Schrödinger equation for vectors, and is expandable in powers of ε. But we saw that we require an extra factor of $e^{-\alpha i/\varepsilon}$ to make the asymptotic series work, immediately giving the contradiction. Not allowing these changes of basis therefore prevents $\tilde{H} = 0$.

4 Recursion on the sphere

This section will focus on solving the recursion given by the asymptotic series on the sphere. We will use it to construct an approximate solution $\mathbf{x}^{(n)}$ to the von Neumann equation, such that $\mathbf{x}^{(n)} \to (0, 0, -1)$ as $t \to -\infty$, that is which starts at the ground state of the Hamiltonian.

4.1 Obtaining the recursion

Consider the von Neumann equation (11) on the sphere and the corresponding conditions on the asymptotic series (14) which, for $\mathbf{x}_n = (x_n, y_n, z_n)$, can be written as

$$\dot{x}_n = -y_{n+1} - \dot{\theta} z_n, \tag{17}$$

$$\dot{y}_n = x_{n+1}, \tag{18}$$

$$\dot{z}_n = \dot{\theta} x_n. \tag{19}$$

Note that (13,15) imply $\mathbf{x}_0 = \pm e_3$, moreover considering a solution which starts at $-e_3$ enforces $\mathbf{x}_0 = -e_3$ and $\mathbf{x}_n \to 0$ for $n \geq 1$ as $t \to -\infty$. So in coordinates $x_0 = y_0 = 0, z_0 = -1$ and all other components x_n, y_n, z_n decay at infinity.

Straightforward calculations show that $x_1 = z_1 = 0$ and $y_1 = \dot{\theta}$, from which it is easy to check that the recursion obtained here and in [BT05a] are completely equivalent. Letting tildes denote the quantities obtained in Proposition 2 of [BT05a], we have $x_n = -2\tilde{z}_n, y_n = 2i\tilde{x}_n$ and $z_n = 2\tilde{y}_n$. Note that Betz and Teufel require more work, because they solve the matrix recursion for the corresponding projection, and only after do they introduce a basis like the Pauli spin matrices[12]. Their approach is easier to generalise, but in this 2-dimensional problem, introducing the Pauli spin matrices right away gives the same recursion with little effort.

Observe that one can use the above recursion to obtain:

$$\begin{aligned} x_{n+2} &= \dot{y}_{n+1} && \text{by (18)}, \\ &= -\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{x}_n + \dot{\theta} z_n\right) && \text{by (17)}, \\ &= -\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{x}_n + \dot{\theta}\int_{-\infty}^{t} \dot{\theta}(s) x_n(s)\mathrm{d}s\right) && \text{by (19)}, \end{aligned}$$

where the last integral has the given bounds for $n \geq 1$, since $z_n \to 0$ as $t \to -\infty$. This is to ensure that x_{n+2} decays at $-\infty$ as well. We therefore obtain the 2-step recursion

$$x_{n+2} = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{x}_n + \dot{\theta}\int_{-\infty}^{t} \dot{\theta}(s) x_n(s)\mathrm{d}s\right), \tag{20}$$

for $n \geq 1$. For $n = 0$, $z_n \to -1$ as $t \to -\infty$ so the integration constant would be different. Instead, calculating x_2 directly using (18) gives $x_2 = \ddot{\theta}$. This recursion allows us to determine all x_n's, with y_n, z_n easily following from the system of ODEs. It is obvious that $x_n = z_n = 0$ for n odd and $y_n = 0$ for n even, so let n be even for this section, and let us focus on (20).

[12] Note that the basis they chose really are the Pauli spin matrices, but expressed them in the adiabatic basis, a consequence of not changing basis from the start.

4.2 Poles of decaying functions

Before analysing this recursion, we go through a complex analysis interlude, to motivate the approach used. Let $g : \mathbb{C} \to \mathbb{C} \cup \{\infty\}$ be a meromorphic function decaying at infinity, that is $g(z) \to 0$ as $|z| \to \infty$. For simplicity, assume that g has a pole at some $w \in \mathbb{C}$ and no other singularities. We can use Laurent's annulus theorem to expand g on $\mathbb{C} \setminus \{w\}$, and write

$$g(z) = \sum_{k=-n}^{\infty} a_k (z-w)^k$$

for some constants $a_k \in \mathbb{C}$, where $n \in \mathbb{N}$ is the order of the pole at w. Let h be another meromorphic function, defined by

$$h(z) = \sum_{k=-n}^{-1} a_k (z-w)^k,$$

where the series converges on $\mathbb{C} \setminus \{w\}$. The function $g - h$ has a removable singularity at w, and is therefore an entire function. Now observe that h decays at infinity and hence so does $g - h$. We therefore have a bounded entire function, and Liouville's theorem tells us that $g - h$ is equal to some constant, which must be 0 since the function decays. We therefore have $g \equiv h$, which gives $a_k = 0$ for $k \geq 0$, meaning that the behaviour of g is completely determined by its pole at w. This argument can easily be repeated for decaying functions with finitely many poles, where we can discard the non-negative powers of z.

The following are then obvious using Laurent series:

- $\frac{\mathrm{d}}{\mathrm{d}z} g$ has the same poles as g, with all orders increased by 1,
- if all poles have order at least 2, $\int g$ has the same poles as g, with all orders decreased by 1,
- gh has poles at the same points as g and h, with the respective orders summed,
- $g+h$ has poles at the same points as g and h, where the order of each is the maximum of the respective orders for g and h.

Let us return to our recursion and recall our special choice of $\dot{\theta}(t) = \frac{2\gamma t_c}{t^2+t_c^2}$, which we will write as

$$\dot{\theta} = \frac{\gamma}{t_c}(f + \overline{f}) = \frac{2\gamma}{t_c} \operatorname{Re}(f), \text{ for } f(t) = \frac{i t_c}{t + i t_c}. \tag{21}$$

This makes clear that $\dot{\theta}$ is meromorphic with a pole of order 1 at each $\pm i t_c$, and that $\dot{\theta}$ decays at infinity. Note that the argument t is real, but we can nonetheless use complex analysis to obtain useful information.

A closer look at (20) shows that the order of each pole increases by 2 at each iteration of

the 2-step recursion, and since we start with $x_2 = \ddot{\theta}$ having poles of order 2, one can write

$$x_n(t) = \sum_{k=1}^{n} \left(\frac{a_k}{(t + it_c)^k} + \frac{b_k}{(t - it_c)^k} \right),$$

for some coefficients a_k, b_k. Note that there are no terms with non-negative exponents of $(t \pm it_c)$ as each x_n decays at infinity.

4.3 Analysis of the recursion

We now use an ansatz, effectively using the Laurent series argument above, but writing with f and $\overline{f}$ as in (21) instead of $1/(t \pm it_c)^k$. This choice is to make the algebra of analysing the recursion easier, and the following result is exactly as Proposition 5 in [BT05a].

Proposition 1. *Let $n \geq 2$ be an even integer, $0 \leq j \leq n-1$ and define the coefficients $a_j^{\{n\}}$ through the following recursion:*

$$\begin{aligned} a_0^{\{2\}} &= 2,\ a_1^{\{2\}} = 0, \\ a_j^{\{n+2\}} &= \frac{n+1-j}{n(n+1)} \left((n-j) a_j^{\{n\}} - \gamma^2 \sum_{k=0}^{j} \frac{1}{n-k} \sum_{m=0}^{k} a_m^{\{n\}} \right) \text{ for } j < n, \\ a_n^{\{n+2\}} &= a_{n-1}^{\{n+2\}}, a_{n+1}^{\{n+2\}} = 0. \end{aligned} \tag{22}$$

Then

$$x_n = \frac{-\gamma(n-1)!}{t_c^n} \sum_{j=0}^{n-1} 2^{-j} a_j^{\{n\}} \operatorname{Im}(f^{n-j}), \tag{23}$$

$$y_{n+1} = \frac{\gamma n!}{t_c^{n+1}} \sum_{j=0}^{n} 2^{-j} \left(\frac{n+1}{n+1-j} a_j^{\{n+2\}} \right) \operatorname{Re}(f^{n+1-j}), \tag{24}$$

$$z_n = \frac{\gamma^2 (n-1)!}{t_c^n} \sum_{j=0}^{n-1} 2^{-j} \left(\frac{1}{n-j} \sum_{k=0}^{j} a_k^{\{n\}} \right) \operatorname{Re}(f^{n-j}). \tag{25}$$

Note that the ordering of the $a_j^{\{n\}}$'s is somewhat mismatched to the powers of f, that is a small index j in $a_j^{\{n\}}$ corresponds to a high power of f. This turns out to be a more natural definition, and reversing this order only complicates things.

Comparing this result to Betz and Teufel's Proposition 5, one notes that the coefficients $a_j^{\{n\}}$ have a different starting point, that is their $a_0^{\{2\}} = 1$, which makes up for the factor of 2 of difference we have between our components x, y, z[13]. Most following results will be similar to that of Betz and Teufel's, often with a difference of a factor of 2.

[13] One can also note the difference in notation, we write $a_k^{\{n\}}$ for their $a_k^{(n)}$, which is because we will use the superscript $^{(n)}$ to mean partial sums later on.

Proof. We will prove by induction on n that (23) holds, only really using algebraic manipulations along with (20). One can then compute the result for y_{n+1} and z_n with (23) and (18,19).

Observe that $\dot{f} = (i/t_c)f^2$ and $f + \overline{f} = 2f\overline{f}$, which inductively yield

$$\dot{\theta}\,\mathrm{Re}(f^{n-j}) = \frac{\gamma}{t_c}\left(2^{-n+j+1}\,\mathrm{Re}(f) + \sum_{k=j}^{n-1} 2^{j-k}\,\mathrm{Re}(f^{n-k+1})\right), \tag{26}$$

$$\dot{\theta}\,\mathrm{Im}(f^{n-j}) = \frac{\gamma}{t_c}\sum_{k=j}^{n-1} 2^{j-k}\,\mathrm{Im}(f^{n-k+1}), \tag{27}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{Re}(f^m) = -\frac{m}{t_c}\,\mathrm{Im}(f^{m+1}), \tag{28}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{Im}(f^m) = \frac{m}{t_c}\,\mathrm{Re}(f^{m+1}), \tag{29}$$

which is exactly the content of lemma 4 in [BT05a], up to a correction in the sign for the derivatives. Noting that $x_2 = \ddot{\theta}$ from our calculation of y_1 with (18), and recalling that $\dot{\theta} = 2\gamma/t_c\,\mathrm{Re}(f)$ immediately yields $x_2 = -2\gamma/t_c^2\,\mathrm{Im}(f^2)$, thus comparing coefficients proves the base case $n = 2$.

Now suppose that x_n is given by (23) for some even n. There are 2 terms in the recursion (20) we need to calculate, and we start with the easier one:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\dot{x}_n &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{-\gamma(n-1)!}{t_c^{n+1}}\sum_{j=0}^{n-1} 2^{-j}a_j^{\{n\}}(n-j)\,\mathrm{Re}(f^{n-j+1}),\\ &= \frac{\gamma(n-1)!}{t_c^{n+2}}\sum_{j=0}^{n-1} 2^{-j}a_j^{\{n\}}(n-j)(n-j+1)\,\mathrm{Im}(f^{n-j+2}),\end{aligned} \tag{30}$$

first using (29) then (28).

We now calculate the second term in (20), which we will do step by step: we have

$$\begin{aligned}\dot{\theta}x_n &= \frac{-\gamma(n-1)!}{t_c^n}\sum_{j=0}^{n-1} 2^{-j}a_j^{\{n\}}\frac{\gamma}{t_c}\sum_{k=j}^{n-1} 2^{j-k}\,\mathrm{Im}(f^{n-k+1}), && \text{using (27),}\\ &= \frac{-\gamma^2(n-1)!}{t_c^{n+1}}\sum_{k=0}^{n-1} 2^{-k}\,\mathrm{Im}(f^{n-k+1})\sum_{j=0}^{k} a_j^{\{n\}}, && \text{switching order of summation.}\end{aligned}$$

We now need to integrate, and at first glance integrating $\mathrm{Im}(f)$ could be problematic, in the sense that we would not be able to write the result as a sum of powers of f and $\overline{f}$, but this is not an issue since all exponents of f are larger than or equal to 2 in the above. So

using (28) and the Fundamental Theorem of Calculus yields

$$\int_{-\infty}^{t} \dot{\theta} x_n = \frac{\gamma^2 (n-1)!}{t_c^n} \sum_{k=0}^{n-1} 2^{-m} b_k^{\{n\}} \operatorname{Re}(f^{n-k}),$$

where we define $b_k^{\{n\}} = \frac{1}{n-k} \sum_{j=0}^{m} a_j^{\{n\}}$ like in [BT05a], to reduce text. Observe that this term exactly occurs in the recursion (22), and that this expression is exactly z_n in (25). Now use (26) to get

$$\begin{aligned}
\dot{\theta} \int_{-\infty}^{t} \dot{\theta} x_n &= \frac{\gamma^2 (n-1!)}{t_c^n} \sum_{k=0}^{n-1} 2^{-k} b_k^{\{n\}} \frac{\gamma}{t_c} \left(2^{-n+k+1} \operatorname{Re}(f) + \sum_{j=k}^{n-1} 2^{k-j} \operatorname{Re}(f^{n-j+1}) \right), \\
&= \frac{\gamma^3 (n-1!)}{t_c^{n+1}} \left(2^{-n+1} \operatorname{Re}(f) \sum_{k=0}^{n-1} b_k^{\{n\}} + \sum_{j=0}^{n-1} 2^{-j} \operatorname{Re}(f^{n-j+1}) \sum_{k=0}^{j} b_k^{\{n\}} \right),
\end{aligned}$$

switching the order of summation for the latter part of the last equality. The only thing left to do is to differentiate, giving

$$\frac{\mathrm{d}}{\mathrm{d}t} \dot{\theta} \int_{-\infty}^{t} \dot{\theta} x_n = \frac{-\gamma^3 (n-1!)}{t_c^{n+2}} \left(2^{-n+1} \operatorname{Im}(f^2) \sum_{k=0}^{n-1} b_k^{\{n\}} + \sum_{j=0}^{n-1} 2^{-j} \operatorname{Im}(f^{n-j+2}) \sum_{k=0}^{j} b_k^{\{n\}} \right). \tag{31}$$

Combining (31) with (30) and using (20), we get

$$\begin{aligned}
x_{n+2} = \frac{-\gamma (n-1)!}{t_c^{n+1}} \Bigg(&- \gamma^2 2^{-n+1} \operatorname{Im}(f^2) \sum_{k=0}^{n-1} b_k^{\{n\}} \\
&+ \sum_{j=0}^{n-1} 2^{-j} (n-j+1) \operatorname{Im}(f^{n-j+1}) \left((n-j) a_j^{\{n\}} - \gamma^2 \sum_{k=0}^{j} b_k^{\{n\}} \right) \Bigg),
\end{aligned}$$

which ends the proof for x_n after comparing coefficients. We have also proved the result for z_n and one could show it for y_{n+1} using (18). This would not be very instructive and follows the same sort of computations, so is omitted. □

We have transferred the analysis of x_n to that of the coefficients $a_k^{\{n\}}$. The next result contains the asymptotic behaviour needed, and is exactly Proposition 6 from [BT05a].

Proposition 2. *For $a_k^{\{n\}}$ defined as in Proposition 1, we have*

(a) $a_0^{\{n\}} = \frac{4 \sin \gamma\pi/2}{\gamma\pi} (1 + O(1/n^2))$;

(b) there is some $C_1 > 0$ such that

$$|a_1^{\{n\}}| \leq C_1 \frac{\log n}{n-1}$$

for all $n \in \mathbb{N}$*;*

(c) for every $p > 1$*, there is some* $C_2 > 0$ *such that*

$$|a_j^{\{n\}}| \leq \frac{C_2 p^j}{n-1}$$

for all $j \geq 2$ *and* $n \in \mathbb{N}$.

Proof. (a) Recall that $a_0^{\{2\}} = 2$ and $a_0^{\{n+2\}} = (1 - \gamma^2/n^2)$, from (22). Now consider the following infinite product

$$\sin(\pi x) = \pi x \prod_{m=1}^{\infty} \left(1 - \frac{x^2}{m^2}\right),$$

which one can find in [AS64], 4.3.89. We have

$$\begin{aligned} \sin(\pi\gamma/2) &= \frac{\pi\gamma}{2} \prod_{m=1}^{\infty} \left(1 - \frac{\gamma^2}{(2m)^2}\right), \\ &= \frac{\gamma\pi}{4} a_0^{\{n\}} \prod_{m=n/2}^{\infty} \left(1 - \frac{\gamma^2}{(2m)^2}\right), \end{aligned}$$

giving the result. In particular, $a_0^{\{n\}}$ is bounded.

(b) Let $\alpha_n = (n-1)a_1^{\{n\}}$, we want to show $|\alpha_n| \leq C_1 \log n$ and proceed by induction, for which the base case of $n = 2$ is obvious. Then

$$\begin{aligned} |\alpha_{n+2}| &= \left|\alpha_n \left(1 - \frac{\gamma^2}{(n-1)^2}\right) - \gamma^2 \left(\frac{1}{n} + \frac{1}{n-1}\right) a_0^{\{n\}}\right|, && \text{by (22)}, \\ &\leq |\alpha_n| \left(1 - \frac{\gamma^2}{(n-1)^2}\right) + \tilde{C} \left(\frac{1}{n} + \frac{1}{n-1}\right), && \text{for some } \tilde{C} > 0 \text{ as } a_0^{\{n\}} \text{ is bounded}, \\ &\leq C_1 \log n + \tilde{C} \left(\frac{1}{n} + \frac{1}{n-1}\right), && \text{taking } n - 1 > \gamma \text{ and by induction}. \end{aligned}$$

Now let $\beta_n \in (0, 2)$ be given by the Mean Value Theorem such that $\log(n+2) - \log n = 2/(n + \beta_n)$, then we require taking C_1 such that

$$\tilde{C} \left(\frac{1}{n} + \frac{1}{n-1}\right) \leq \frac{2C_1}{n + \beta_n},$$

which can be done uniformly in n, and we have shown $|\alpha_{n+2}| \leq C_1 \log(n+2)$. So (b) is done, and in particular $a_1^{\{n\}}$ is bounded.

(c) Let $p > 1$, define $c_j^{\{n\}} = (n-1)p^{-j} a_j^{\{n\}}$ and let $c^{\{n\}} = \max_{j \geq 2} |c_j^{\{n\}}|$. We show that $c^{\{n\}}$ is bounded uniformly in n. Without loss of generality consider $j \leq n-1$, since $a_n^{\{n\}} = a_{n-1}^{\{n\}}$ so up to a multiple of p the bound is the same.

The starting point is using the recursion (22), giving

$$\begin{aligned} c_j^{\{n+2\}} &= p^{-j}\frac{n+1-j}{n}\Bigg[(n-j)a_j^{\{n\}} - \gamma^2\sum_{k=2}^{j}\frac{1}{n-k}\sum_{m=2}^{k}a_m^{\{n\}} \\ &\quad - \gamma^2\left(a_0^{\{n\}}\sum_{k=0}^{j}\frac{1}{n-k} + a_1^{\{n\}}\sum_{k=1}^{j}\frac{1}{n-k}\right)\Bigg], \\ &= \frac{n+1-j}{n(n-1)}\left((n-j)c_j^{\{n\}} - \gamma^2 p^{-j}\sum_{k=2}^{j}\frac{1}{n-k}\sum_{m=2}^{k}p^m c_m^{\{m\}}\right) \\ &\quad - \frac{n+1-j}{n}\gamma^2\left(a_0^{\{n\}}p^{-j}\sum_{k=0}^{j}\frac{1}{n-k} + a_1^{\{n\}}p^{-j}\sum_{k=1}^{j}\frac{1}{n-k}\right). \end{aligned}$$

We now look more closely at each term, starting with

$$p^{-j}\sum_{k=2}^{j}\frac{1}{n-k}\sum_{m=2}^{k}p^m \leq \frac{p^{-j}}{n-j}\sum_{k=2}^{j}p^2\cdot\frac{p^{k-1}-1}{p-1} \leq \frac{p^2}{(n-j)(p-1)^2},$$

where we used $n-j \leq n-k$, the geometric sum formula and discarding the -1 it gives. We also have

$$p^{-j}\sum_{k=0}^{j}\frac{1}{n-k} \leq \frac{(j+1)p^{-j}}{n-j} \leq \frac{1}{(n-j)\log p},$$

counting the number of elements in the sum and taking the maximum for the first inequality, and the second can be shown by proving $(j+1)p^{-j}\log p$ is positive and bounded by 1 for $p > 1$. Here, we also use $p > 1$ to divide the corresponding inequality by $\log p$, a positive number. This same bound holds when starting at $k = 1$, so we bounded both terms in the second brackets. Putting these together gives

$$\begin{aligned} |c_j^{\{n+2\}}| &\leq c^{\{n\}}\left(\frac{(n+1-j)(n-j)}{n(n-1)} + \frac{n+1-j}{n(n-1)}\frac{\gamma^2 p^2}{(n-j)(p-1)^2}\right) \\ &\quad + \frac{n+1-j}{n}\gamma^2(|a_0^{\{n\}}| + |a_1^{\{n\}}|)\frac{1}{(n-j)\log p}. \end{aligned}$$

Using that both $a_0^{\{n\}}, a_1^{\{n\}}$ are bounded, and taking maximum over j gives constants B_1, B_2 such that

$$c^{\{n+2\}} \leq c^{\{n\}}\left(\frac{n-2}{n} + \frac{B_1}{n(n-1)}\right) + \frac{B_2}{n} = c^{\{n\}} + \frac{1}{n}\left(c^{\{n\}}\left(-2 + \frac{B_1}{n-1}\right) + B_2\right),$$

where we used that $(n+1-j)/(n-j)$ is bounded uniformly for $n \geq 2, 2 \leq j \leq n-1$, and that $(n+1-j)(n-j) \leq (n-1)(n-2)$ for the inequality. We can take n large so that

$n-1 > B_2$, then $-2 + B_2/(n-1) < -1$, yielding

$$c^{\{n+2\}} < c^{\{n\}} + \frac{1}{n}(-c^{\{n\}} + B_2).$$

If $c^{\{n\}} \geq B_2$, this gives $c^{\{n+2\}} < c^{\{n\}}$ and if $c^{\{n\}} < B_2$, then $c^{\{n+2\}} < B_2(1 + 1/n)$. So $c^{\{n\}}$ is bounded and (c) is done. □

These directly imply that for all $j \geq 1, \alpha \in (0,1)$,

$$|a_j^{\{n\}}| \leq p^j \frac{C}{(n-1)^\alpha}. \tag{32}$$

We are now ready to state and prove the main theorem of this section, where we combine the previous 2 propositions to get the behaviour of y_{n+1} for large n. This is Theorem 3 in [BT05a], but we omit the result for x_n, z_n as we would not be using them.

Theorem 1. *For n even and $\alpha \in (0,1)$, we have*

$$y_{n+1}(t) = \frac{n!}{t_c^{n+1}} \left(\frac{4\sin(\gamma\pi/2)}{\pi} \operatorname{Re}\left(f^{n+1}(t)\right) + O\left(R_{n+1}^\alpha(t)\right) \right), \tag{33}$$

where the error is defined by

$$R_{n+1}^\alpha(t) = \frac{1}{(n+1)^\alpha} \max\left\{ |f(t)|^{n+1}, \left(\frac{1}{\sqrt{2}}\right)^{n-1} |f(t)|^2 \right\}. \tag{34}$$

For 2 sequences of functions g_n, G_n on $\mathbb{R}$, we write $g_n(t) = O(G_n(t))$ if there exists $C > 0$ such that $|g_n(t)| \leq |G_n(t)|$ for all $t \in \mathbb{R}$ and all $n \in \mathbb{N}$.

Looking at the error, we see that for $|t| \leq t_c$, $|f| \geq 1/\sqrt{2}$ so the first term is maximal, recalling that $f = it_c/(t + it_c)$. Similarly, for $|t| > t_c$, $|f| < 1/\sqrt{2}$ so the second term is maximal. Note that these grow like $n!/t_c^n$, so the asymptotic series for y indeed diverges, as mentioned in the start of section 3.3.

Proof. We use (24) and first consider the $j = 0$ term with the help of proposition 2 (a), which yields

$$\frac{\gamma n!}{t_c^n} a_0^{\{n+2\}} \operatorname{Re}(f^{n+1}) = \frac{n!}{t_c^n} \frac{4\sin(\gamma\pi/2)}{\pi} \left(\operatorname{Re}\left(f^{n+1}\right) + O\left(\frac{1}{n^2}|f|^{n+1}\right) \right)$$

the first part of which gives the first term in (33), and the second part goes in the error as $1/n^2 = O((n+1)^{-\alpha})$.

We are left with showing that all other terms go in the error, that is

$$\sum_{j=1}^{n} 2^{-j} \frac{n+1}{n+1-j} a_j^{\{n+2\}} \operatorname{Re}(f^{n+1-j}(t)) = O(R_{n+1}^{\alpha}(t)). \tag{35}$$

Let us bound the left hand side (LHS) of (35) using that $|\operatorname{Re}(f)| \leq |f|$ and (32), giving

$$|\text{LHS}| \leq \frac{C}{(n+1)^{\alpha}} \sum_{j=1}^{n} \frac{n+1}{n+1-j} \left(\frac{p}{2}\right)^{j} |f|^{n+1-j}.$$

We have 2 cases to consider, that is $|t| \leq t_c$ or $|t| > t_c$. For the former, observe that $|f|^{-1} \leq \sqrt{2}$, which gives

$$|\text{LHS}| \leq \frac{C}{(n+1)^{\alpha}} |f|^{n+1} \sum_{j=1}^{n} \frac{n+1}{n+1-j} \left(\frac{p}{\sqrt{2}}\right)^{j},$$

and for the latter, we have $|f| \leq 1/\sqrt{2}$, which gives

$$|\text{LHS}| \leq \frac{C}{(n+1)^{\alpha}} \left(\frac{1}{\sqrt{2}}\right)^{n-1} |f|^{2} \sum_{j=1}^{n} \frac{n+1}{n+1-j} \left(\frac{p}{\sqrt{2}}\right)^{j}.$$

Comparing with R_{n+1}^{α} in (34), we see that we require that the sum $\sum_{j=1}^{n} \frac{n+1}{n+1-j} \left(p/\sqrt{2}\right)^{j}$ is bounded uniformly in n. The $n+1$ term in the numerator proves problematic, so we use that $(n+1)/(n+1-j) \leq j$ for $2 \leq j \leq n-1$[14]. Then for the terms with $2 \leq j \leq n-1$, the sum

$$\sum_{j=2}^{n-1} \frac{n+1}{n+1-j} \left(\frac{p}{\sqrt{2}}\right)^{j} \leq \sum_{j=2}^{n-1} j \left(\frac{p}{\sqrt{2}}\right)^{j}$$

is uniformly bounded in n if we take $p \in (1, \sqrt{2})$. The $j = 1$ term is $\frac{n+1}{n}(p/\sqrt{2})$ and the $j = n$ term is $(n+1)(p/\sqrt{2})^{n}$, both uniformly bounded. □

5 Superadiabatic basis

Let us go back to the von Neumann equation (5). Recall that one solution allows diagonalising the Hamiltonian, and similarly, having an approximate solution will allow us to render H "almost diagonal". The approximate solution in question will come from the asymptotic series, and will be made precise by the 2 previous sections.

[14] One can prove this by checking that the interval $[2, n-1]$ is always between the roots of $j^2 - (n+1)j + (n+1)$.

For a fixed $n \in \mathbb{N}$, we define the truncated sums

$$x^{(n)} = \sum_{k=0}^{n} \varepsilon^k x_k,\ y^{(n)} = \sum_{k=0}^{n} \varepsilon^k y_k,\ z^{(n)} = \sum_{k=0}^{n} \varepsilon^k z_k.$$

Similarly define $\pi^{(n)} = 1/2(I_2 + x^{(n)}\sigma_x + y^{(n)}\sigma_y + z^{(n)}\sigma_z)$. Each component behaves like $\varepsilon^n n!/t_c^n$, so these series will not converge as $n \to \infty$. This motivates the idea of choosing n not to be fixed, but instead as a function of ε, and for it to be optimal for the given ε.

In the first chapter of [Din73], one encounters an asymptotic series where the terms in the expansion first decrease in magnitude, then reach a minimum and increase again, causing the divergence. The optimal place to truncate here is then the minimal term, and this is exactly our case here[15]. Up to some constants, we therefore want to minimise an expression of the form $\varepsilon^n n!/t_c^n$ as a function of n, which is equivalent to minimising $\sqrt{2\pi n}(n\varepsilon/et_c)^n$ by Stirling's formula. Taking logarithms, and differentiating gives $1/2n + \log(n\varepsilon/t_c) = 0$. For n large, this has approximate solution $n \approx t_c/\varepsilon$, which for ε small is indeed large. However, this value will most likely not be an integer, so we require a slight adjustment. Just as in [BT05a], define

$$n_\varepsilon = \frac{t_c}{\varepsilon} - 1 + \tau_\varepsilon, \tag{36}$$

where $\tau_\varepsilon \in [0, 2)$ is such that n_ε is an even integer. Note that the computations needed will usually contain $n_\varepsilon + 1$, which is why we have a -1 in the definition. With this choice, we have $x_{n_\varepsilon+1} = z_{n_\varepsilon+1} = 0$, which will be helpful later.

Let $\pi^{(n_\varepsilon)}$ be the truncated sum. It is a Hermitian matrix and it almost solves the von Neumann equation (5). As it is Hermitian, we can let $U_\varepsilon^{n_\varepsilon}$ be a special unitary matrix which diagonalises $\pi^{(n_\varepsilon)}$, where we write

$$U_\varepsilon^{n_\varepsilon} = \begin{pmatrix} u_1 & -\overline{u}_2 \\ u_2 & \overline{u}_1 \end{pmatrix}, \tag{37}$$

for $|u_1|^2 + |u_2|^2 = 1$. This choice is not unique, but enforcing $\mathbb{R} \ni u_1 \geq 0$ gives uniqueness and will simplify calculations. Then

$$U_\varepsilon^{n_\varepsilon *} \pi^{(n_\varepsilon)} U_\varepsilon^{n_\varepsilon} = \Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \tag{38}$$

for some $\lambda_i \in \mathbb{R}$. This unitary $U_\varepsilon^{n_\varepsilon}$ will then give us a new Hamiltonian through (6), which will have exponentially small off-diagonal elements. The new basis to which $U_\varepsilon^{n_\varepsilon}$ sends us is called the superadiabatic basis, which gives a better representation of the system, since it follows the time evolution only from an error of order ε^n away. We have obtained it in a

[15] Another way of viewing it is with Theorem 2, where one has the off-diagonal elements of the Hamiltonian exactly corresponding to terms of order $n + 1$, thus making it clearer that one wants to minimise the magnitude of a high order term in the expansion.

similar way to how Berry constructs his superadiabatic bases in [Ber90], while in [BT05a] it is obtained through successively diagonalising the Hamiltonian, both of which yield the same result.

Before stating or proving anything about this change of basis, we will require a preliminary lemma, involving the error for $\mathbf{x}^{(n_\varepsilon)}$ to be on the unit sphere, or equivalently the error for $\pi^{(n_\varepsilon)}$ to be a projection.

Lemma 1. *The functions $x^{(n_\varepsilon)}, y^{(n_\varepsilon)}, z^{(n_\varepsilon)}+1$ and their derivatives $\dot{x}^{(n_\varepsilon)}, \dot{y}^{(n_\varepsilon)}, \dot{z}^{(n_\varepsilon)}$ are all $O(\varepsilon\dot{\theta})$.*

This statement is really saying that $x^{(n_\varepsilon)}-x_0, y^{(n_\varepsilon)}-y_0, z^{(n_\varepsilon)}-z_0$ and their derivatives are bounded, and $(x_0,y_0,z_0)=(0,0,-1)$ as mentioned in section 4.1. Note that these are exactly the terms in the truncated sum with a power of ε greater than or equal to 1.

Proof. We first show that there exists a constant $C>0$, independent of k such that each x_k, y_k, z_k are bounded in norm by $C\dot{\theta}(k-1)!/t_c^k$ for all k. We prove it for z_k, and the others are similar. Let z_k be as in (25), from which it is clear we require

$$\sum_{j=0}^{k-1} 2^{-j}\left(\frac{1}{k-j}\sum_{m=0}^{j} a_m^{\{k\}}\right)\mathrm{Re}(f^{k-j}) = O(\dot{\theta}).$$

Now observe that $\mathrm{Re}(f)=(t_c/2\gamma)\dot{\theta}=O(\dot{\theta})$ trivially, and for $l\geq 2$, $|\mathrm{Re}(f^l)|\leq |f|^2|f|^{l-2}=O(\dot{\theta})$, as the first factor is $(t_c/2\gamma)\dot{\theta}$ and the second factor is bounded uniformly in t[16]. The proof is now reduced to showing that the sum above, without $\mathrm{Re}(f^{k-j})$, is bounded uniformly in n. This is obvious for the terms with $j=0,1$ as $a_j^{\{n\}}$ is bounded from Proposition 2 (a), (b), so we are left with bounding

$$\begin{aligned}\left|\sum_{j=2}^{k-1} 2^{-j}\frac{1}{k-j}\sum_{m=2}^{j} a_m^{\{k\}}\right| &\leq \frac{C_2}{k-1}\sum_{j=2}^{k-1}\frac{2^{-j}}{k-j}\sum_{m=2}^{j} p^m, && \text{from Prop 2 (c),}\\ &\leq \frac{C_2 p}{(k-1)(p-1)}\sum_{j=2}^{k-1}\frac{(p/2)^j}{k-j} && \text{using geometric sums,}\end{aligned}$$

which is indeed uniformly bounded in k if we take $p\in(1,2)$. This proves that $|z_k|\leq C\dot{\theta}(k-1)!/t_c^k$ for all k, and the proofs for x_k, y_k are similar. We now consider the truncated sum $|z^{(n_\varepsilon)}+1|$, which is

$$\left|\sum_{k=1}^{n_\varepsilon}\varepsilon^k z_k\right| \leq C\varepsilon\dot{\theta}\sum_{k=0}^{n_\varepsilon-1}\varepsilon^k k!/t_c^{k+1} = \frac{C\varepsilon\dot{\theta}}{t_c}\sum_{k=0}^{n_\varepsilon-1} k!/(n_\varepsilon+1-\tau_\varepsilon)^k,$$

[16]This bound can be used for the statement about y_k, but x_k is written with $\mathrm{Im}(f^l)$ instead. Just as for the real part, if $l\geq 2$ then $|\mathrm{Im}(f^l)|\leq |f|^2|f|^{l-2}=O(\dot{\theta})$, but a possible problem occurs for $l=1$, as $\mathrm{Im}(f)=tt_c/(t^2+t_c^2)\neq O(\dot{\theta})$. However, the coefficient in front of $\mathrm{Im}(f)$ contains $a_{n+1}^{\{n+2\}}=0$, so this is not an issue.

where we substituted the definition of n_ε. This proves the result of $z^{(n_\varepsilon)}+1$ as the last sum is uniformly bounded in n_ε, and the same argument proves the result for $x^{(n_\varepsilon)}, y^{(n_\varepsilon)}$.

One then uses (18) to get $|\dot{y}_k| \leq C\dot{\theta}k!/t_c^{k+1}$ for all k, to which one applies the same argument as above to show the result for $\dot{y}^{(n_\varepsilon)}$, and a similar argument for $\dot{x}^{(n_\varepsilon)}, \dot{z}^{(n_\varepsilon)}$. □

We now use this to get a full understanding of the unitary $U_\varepsilon^{n_\varepsilon}$ and the diagonal Λ, which is the content of the following lemma.

Lemma 2. *We have*

$$\lambda_1 = O(\varepsilon\dot{\theta}),\ \dot{\lambda}_1 = O(\varepsilon\dot{\theta}),\ \lambda_2 = 1 + O(\varepsilon\dot{\theta}),\ \dot{\lambda}_2 = O(\varepsilon\dot{\theta}),$$

as well as

$$u_1 = 1 + O(\varepsilon\dot{\theta}),\ \dot{u}_1 = O(\varepsilon\dot{\theta}),\ u_2 = O(\varepsilon\dot{\theta}),\ \dot{u}_2 = O(\varepsilon\dot{\theta}).$$

The notation is suppressed here, but λ_i, v_i all depend on $\varepsilon, n_\varepsilon$ and t. Here, Big O notation is slightly different to its earlier use: for 2 functions g, G of ε and $t \in \mathbb{R}$, write $g(\varepsilon,t) = O(G(\varepsilon,t))$ if there exist $C>0$ and $\varepsilon_0 > 0$ such that $|g(\varepsilon,t)| \leq C|G(\varepsilon,t)|$ for all $t \in \mathbb{R}, \varepsilon \in (0,\varepsilon_0)$. Recall that $\dot{\theta} \to 0$ as $|t| \to \infty$, so λ_2, u_1 go to 1, while λ_1, u_2 and all derivatives go to 0. Looking at the unitary, it means $U_\varepsilon^{n_\varepsilon} \to I_2$, with $\dot{U}_\varepsilon^{n_\varepsilon} \to 0$.

Proof. Observe that we have the usual relations

$$\lambda_1\lambda_2 = \det \pi^{(n_\varepsilon)},\ \lambda_1 + \lambda_2 = \operatorname{Tr} \pi^{(n_\varepsilon)},$$

which immediately give

$$\lambda_1 = \frac{1}{2}\left(1 - \sqrt{1 - 4\det \pi^{(n_\varepsilon)}}\right),\ \lambda_2 = 1 - \lambda_1, \tag{39}$$

since $\operatorname{Tr} \pi^{(n_\varepsilon)} = 1$ thanks to our use of the Pauli spin matrices. Recall that $\det \pi = 1/4(1-(x^2+y^2+z^2))$ for a projection π written in the Pauli basis, thus yielding

$$\begin{aligned} 1 - 4\det \pi^{(n_\varepsilon)} &= (x^{(n_\varepsilon)})^2 + (y^{(n_\varepsilon)})^2 + (z^{(n_\varepsilon)} + 1 - 1)^2, \\ &= O(\varepsilon\dot{\theta})^2 - 2O(\varepsilon\dot{\theta}) + 1, \\ &= 1 + O(\varepsilon\dot{\theta}), \end{aligned}$$

where we used lemma 1 to bound the components x, y, z. We then use the Taylor expansion of $\sqrt{1+s}$ near $s=0$ to get $\sqrt{1 - 4\det \pi^{(n_\varepsilon)}} = 1 + O(\varepsilon\dot{\theta})$, which gives the result for λ_1, from which the result for λ_2 immediately follows.

One can differentiate λ_1 in (39) and use the bounds on the derivatives in lemma 1, with similar arguments, to get the result for $\dot{\lambda}_1$ and $\dot{\lambda}_2$.

To prove the statements involving the components of the unitary matrix, note that its columns $u = \binom{u_1}{u_2}, u^\perp = \binom{-\overline{u}_2}{u_1}$ are exactly the eigenvectors of $\pi^{(n_\varepsilon)}$, and we can write $\pi^{(n_\varepsilon)} = \lambda_1 u u^* + (1-\lambda_1) u^\perp u^{\perp *}$, where $v^* = (\overline{v}_1 \ \overline{v}_2)$ is the covector dual to $v = \binom{v_1}{v_2}$. Expanding this, equating it to $1/2(I_2 + x^{(n_\varepsilon)}\sigma_x + y^{(n_\varepsilon)}\sigma_y + z^{(n_\varepsilon)}\sigma_z)$, and comparing the $(1,1)$ and $(2,1)$ entries of both matrices yields

$$u_1^2(2\lambda_1 - 1) + 1 - \lambda_1 = \frac{1}{2}\left(1 + z^{(n_\varepsilon)}\right), \tag{40}$$

$$u_1 u_2 (2\lambda_1 - 1) = \frac{1}{2}\left(x^{(n_\varepsilon)} + i y^{(n_\varepsilon)}\right). \tag{41}$$

We can rearrange (40) to

$$u_1^2 = \frac{1/2(1 + z^{(n_\varepsilon)}) + \lambda_1 - 1}{2\lambda_1 - 1},$$

which gives that $v_1^2 = 1 + O(\varepsilon\dot{\theta})$, and we use the Taylor expansion of the square root near 1 to get $v_1 = 1 + O(\varepsilon\dot{\theta})$. Similarly, rearrange (41) to

$$u_2 = \frac{x^{(n_\varepsilon)} + i y^{(n_\varepsilon)}}{2u_1(2\lambda_1 - 1)},$$

giving the result for u_2. Differentiating the above easily gives the result for $\dot{u}_1$ and $\dot{u}_2$. □

We can now express the Hamiltonian in this new basis using (6), which will be a crucial intermediate step towards our main Theorem, to render the off-diagonal elements exponentially small.

Theorem 2. *For $U_\varepsilon^{n_\varepsilon}$ as in (37), the new Hamiltonian $H_\varepsilon^{n_\varepsilon} = U_\varepsilon^{n_\varepsilon *} H U_\varepsilon^{n_\varepsilon} - i\varepsilon U_\varepsilon^{n_\varepsilon *}\dot{U}_\varepsilon^{n_\varepsilon}$ is*

$$H_\varepsilon^{n_\varepsilon} = \frac{1}{2}\begin{pmatrix} 1 + O(\varepsilon^2\dot{\theta}) & \varepsilon^{n_\varepsilon+1}(-x_{n_\varepsilon+1} + i y_{n_\varepsilon+1})(1 + O(\varepsilon\dot{\theta})) \\ \varepsilon^{n_\varepsilon+1}(-x_{n_\varepsilon+1} - i y_{n_\varepsilon+1})(1 + O(\varepsilon\dot{\theta})) & -1 + O(\varepsilon^2\dot{\theta}) \end{pmatrix}. \tag{42}$$

Note that our choice of n_ε implies $x_{n_\varepsilon+1} = 0$, thus only the asymptotic of y need to be understood and Theorem 1 will be enough to determine the off-diagonal elements. This is another justification of why we require making the $(n_\varepsilon + 1)$-st term in the truncated sum as small as possible, precisely because it appears in the off-diagonal elements of the Hamiltonian.

Proof. $H_\varepsilon^{n_\varepsilon}$ is still traceless and Hermitian, so we will only need to determine the top row, that is the $(1,1)$-entry h_{11} and the $(1,2)$-entry h_{12}. An explicit calculation with H as in (8) gives

$$h_{11} = 1/2 - u_2\overline{u}_2 - i\varepsilon\left(u_1\dot{u}_1 + \overline{u}_2\dot{u}_2 + \dot{\theta}\,\mathrm{Re}(u_1 u_2)\right),$$

for which one can use Lemma 2 to get h_{11} as in (42). Using this method gives $h_{12} = O(\varepsilon\dot{\theta})$, which is nothing new as it is already the case for the Hamiltonian in the adiabatic basis,

motivating another approach.

Instead, we compute how close $\pi^{(n_\varepsilon)}$ is to solving the von Neumann equation using the recursions (17) - (19), that is

$$i\varepsilon\dot{\pi}^{(n_\varepsilon)} - [H, \pi^{(n_\varepsilon)}] = i\varepsilon^{n_\varepsilon+1}/2(x_{n_\varepsilon+1}\sigma_y - y_{n_\varepsilon+1}\sigma_x). \tag{43}$$

We now use that $U_\varepsilon^{n_\varepsilon}$ is a unitary matrix that diagonalises $\pi^{(n_\varepsilon)}$, so it exactly satisfies $U_\varepsilon^{n_\varepsilon *}U_\varepsilon^{n_\varepsilon} = I_2$ and (38). Differentiating both yields $\dot{U}_\varepsilon^{n_\varepsilon *}U_\varepsilon^{n_\varepsilon} + U_\varepsilon^{n_\varepsilon *}\dot{U}_\varepsilon^{n_\varepsilon} = 0$ and $U_\varepsilon^{n_\varepsilon *}\dot{\pi}^{(n_\varepsilon)}U_\varepsilon^{n_\varepsilon} = \dot{\Lambda} - \dot{U}_\varepsilon^{n_\varepsilon *}\pi^{(n_\varepsilon)}U_\varepsilon^{n_\varepsilon} - U_\varepsilon^{n_\varepsilon *}\pi^{(n_\varepsilon)}\dot{U}_\varepsilon^{n_\varepsilon}$, giving

$$U_\varepsilon^{n_\varepsilon *}\left(i\varepsilon\dot{\pi}^{(n)} - [H, \pi^{(n)}]\right)U_\varepsilon^{n_\varepsilon} = i\varepsilon\dot{\Lambda} - [\tilde{H}, \Lambda]. \tag{44}$$

We can explicitly compute $U_\varepsilon^{n_\varepsilon *}\sigma_x U_\varepsilon^{n_\varepsilon}$ and $U_\varepsilon^{n_\varepsilon *}\sigma_y U_\varepsilon^{n_\varepsilon}$ to find

$$U_\varepsilon^{n_\varepsilon *}\sigma_x U_\varepsilon^{n_\varepsilon} = \begin{pmatrix} u_1(u_2+\overline{u}_2) & u_1^2 - \overline{u}_2^2 \\ u_1^2 - u_2^2 & -u_1(u_2+\overline{u}_2) \end{pmatrix},\ U_\varepsilon^{n_\varepsilon *}\sigma_y U_\varepsilon^{n_\varepsilon} = i\begin{pmatrix} u_1(\overline{u}_2-u_2) & -u_1^2 - \overline{u}_2^2 \\ u_1^2 + u_2^2 & u_1(u_2-\overline{u}_2) \end{pmatrix}.$$

Then compare the $(1,2)$-th entries[17] on the left and right hand sides of (44), using (43) to obtain

$$(\lambda_1 - \lambda_2)h_{12} = \varepsilon^{n+1}/2\left(x_{n_\varepsilon+1}(u_1^2 + \overline{u}_2^2) + iy_{n_\varepsilon+1}(\overline{u}_2^2 - u_1^2)\right)).$$

Lemma 2 then gives h_{12} as in (42). □

6 Solving the equation

6.1 New Hamiltonian

Combining Theorems 1 and 2 allows us to change the Hamiltonian into a form with exponentially small off-diagonal elements, corresponding exactly to Theorem 1 in [BT05a]. But before the statement, we require a rather technical analytic lemma, the proof of which can be found in [BT05a], cf Lemma 5.

Lemma 3. *Uniformly in $x \in [0,1]$ and for $k > 0$, we have*

$$(1+x)^{-k} = e^{-kx} + e^{-kx/2}O(1/k).$$

We can take $x = a/k$ to get $(1+a/k)^{-k} = e^{-a}(1 + e^{a/2}O(1/k))$, which gives

$$(1+a/k)^{-k} = e^{-a}(1 + O(1/k)) \tag{45}$$

uniformly on compact intervals of a.

[17] This method would not provide any information on the diagonal entries of $\tilde{H}$, as the commutator term would cancel them.

Theorem 3. *If we let* $H_\varepsilon^{n_\varepsilon} = U_\varepsilon^{n_\varepsilon *} H U_\varepsilon^{n_\varepsilon} - i\varepsilon U_\varepsilon^{n_\varepsilon *} \dot{U}_\varepsilon^{n_\varepsilon}$*, then we can write*

$$H_\varepsilon^{n_\varepsilon} = \begin{pmatrix} \frac{1}{2} + O(\varepsilon^2\dot{\theta}) & c_\varepsilon^{n_\varepsilon}(t) \\ \overline{c}_\varepsilon^{n_\varepsilon}(t) & -\frac{1}{2} + O(\varepsilon^2\dot{\theta}) \end{pmatrix},$$

where for every $\alpha \in (0,1)$*,*

$$c_\varepsilon^{n_\varepsilon}(t) = 2i\sqrt{\frac{2\varepsilon}{\pi t_c}} \sin\left(\frac{\pi\gamma}{2}\right) e^{-\frac{t_c}{\varepsilon}} e^{-\frac{t^2}{2\varepsilon t_c}} \cos\left(\frac{t}{\varepsilon} - \frac{t^3}{3\varepsilon t_c^2} + \frac{\tau_\varepsilon t}{t_c}\right) + O(\sqrt{\varepsilon} e^{-\frac{t_c}{\varepsilon}} \phi^\alpha(\varepsilon, t)) \quad (46)$$

and

$$\phi^\alpha(\varepsilon, t) = \begin{cases} \varepsilon^\alpha \exp\left(-\frac{t^2}{4\varepsilon t_c}\right) & \text{if } |t| \le t_c, \\ \frac{1}{1+t^2} \exp\left(-\frac{t_c \log 2}{2\varepsilon}\right) & \text{if } |t| > t_c. \end{cases} \quad (47)$$

The content of Theorem 2 was that in a superadiabatic basis of order n, the off-diagonal elements of the Hamiltonian become of order $n+1$. But the crucial point here is that if we choose this n optimally as a function of ε, we can have exponentially small off-diagonal elements, a step entirely out of reach of superadiabatic theory.

Proof. Theorem 2 already gives the diagonal elements and $c_\varepsilon^{n_\varepsilon} = \frac{i}{2}\varepsilon^{n_\varepsilon+1} y_{n_\varepsilon+1}(1 + O(\varepsilon\dot{\theta}))$, as $x_{n_\varepsilon+1} = 0$. Then applying Theorem 1 yields

$$c_\varepsilon^{n_\varepsilon} = i\frac{\varepsilon^{n_\varepsilon+1} n_\varepsilon!}{t_c^{n_\varepsilon+1}} \left(\frac{2\sin(\gamma\pi/2)}{\pi} \operatorname{Re}(f^{n_\varepsilon+1}) + O(R_{n_\varepsilon+1}^\alpha)\right). \quad (48)$$

The factor $\varepsilon^{n_\varepsilon+1} n_\varepsilon! / t_c^{n_\varepsilon+1}$ can be computed using the optimality of n_ε, as in lemma 6 of [BT05a]. Stirling's formula for $n_\varepsilon + 1$ gives

$$n_\varepsilon! = \sqrt{\frac{2\pi}{n_\varepsilon+1}} e^{-(n_\varepsilon+1)} \left(\frac{1}{n_\varepsilon+1}\right)^{-(n_\varepsilon+1)} \left(1 + O\left(\frac{1}{n_\varepsilon+1}\right)\right),$$

and using (36) yields

$$\varepsilon^{n_\varepsilon+1} = t_c^{n_\varepsilon+1} (n_\varepsilon + 1 - \tau_\varepsilon)^{-(n_\varepsilon+1)}.$$

Combining these 2 gives

$$\begin{aligned} \frac{\varepsilon^{n_\varepsilon+1} n_\varepsilon!}{t_c^{n_\varepsilon+1}} &= \sqrt{\frac{2\pi}{n_\varepsilon+1}} e^{-(n_\varepsilon+1)} \left(1 - \frac{\tau_\varepsilon}{n_\varepsilon+1}\right)^{-(n_\varepsilon+1)} \left(1 + O\left(\frac{1}{n_\varepsilon+1}\right)\right), \\ &= \sqrt{\frac{2\pi}{n_\varepsilon+1}} e^{-(n_\varepsilon+1-\tau_\varepsilon)} \left(1 + O\left(\frac{1}{n_\varepsilon+1}\right)\right), \\ &= \sqrt{\frac{2\pi}{n_\varepsilon+1}} e^{-t_c/\varepsilon} (1 + O(\varepsilon)), \\ &= \sqrt{\frac{2\pi\varepsilon}{t_c}} e^{-t_c/\varepsilon} (1 + O(\varepsilon)), \end{aligned}$$

where we used (45) for $a = -\tau_\varepsilon, k = n_\varepsilon + 1$ for the second equality, the definition of n_ε for the third, and the last equality holds since

$$\sqrt{\frac{1}{n_\varepsilon + 1}} = \sqrt{\frac{\varepsilon}{t_c + \varepsilon\tau_\varepsilon}} = \sqrt{\frac{\varepsilon}{t_c}}(1 + O(\varepsilon)).$$

Putting this with (48) yields

$$c_\varepsilon^{n_\varepsilon} = 2i\sqrt{\frac{2\varepsilon}{\pi t_c}} e^{-t_c/\varepsilon} \left(\sin(\gamma\pi/2)\,\mathrm{Re}(f^{n_\varepsilon+1}) + O(R^\alpha_{n_\varepsilon+1})\right), \tag{49}$$

so comparing with (46), we have reduced the proof to showing that

$$\mathrm{Re}(f^{n_\varepsilon+1}(t)) + O(R^\alpha_{n_\varepsilon+1}(t)) = e^{-\frac{t^2}{2\varepsilon t_c}} \cos\left(\frac{t}{\varepsilon} - \frac{t^3}{3\varepsilon t_c^2} + \frac{\tau_\varepsilon t}{t_c}\right) + O(\phi^\alpha(\varepsilon, t)). \tag{50}$$

For $|t| > t_c$, the leading order term involving cos is $O(\phi^\alpha(\varepsilon, t))$, so we check that both terms on the left hand side of (50) are $O(\phi^\alpha(\varepsilon, t))$. We have

$$\begin{aligned}
|\mathrm{Re}(f^{n_\varepsilon+1})| \le |f|^{n_\varepsilon+1} &= \left(1 + \frac{t^2}{t_c^2}\right)^{-1} |f|^{n_\varepsilon-1}, \\
&\le \left(1 + \frac{t^2}{t_c^2}\right)^{-1} 2^{-(n_\varepsilon-1)/2}, && \text{as } |f| \le 1/\sqrt{2} \text{ for } |t| \ge t_c, \\
&= \left(1 + \frac{t^2}{t_c^2}\right)^{-1} \exp\left(\frac{\log 2}{2}(-t_c/\varepsilon + 2 - \tau_\varepsilon)\right), && \text{using (36)}, \\
&= O(\phi^\alpha(\varepsilon, t)).
\end{aligned}$$

The analysis for $R^\alpha_{n_\varepsilon+1}$ is the same, with an extra factor of $(n_\varepsilon + 1)^{-\alpha}$, which is $O(\varepsilon^\alpha) = O(1)$ so can be discarded.

For $|t| \le t_c$, we will determine the leading order term, thus requiring analysis of the modulus as well as the phase. We need to split the range of times further, as the error will still dominate the leading order term in a certain range of t. More precisely, we consider $|t| \ge \varepsilon^{\alpha/2}$ or $|t| < \varepsilon^{\alpha/2}$.

Observe that

$$|f|^{n_\varepsilon+1} = \left(1 + \frac{t^2}{t_c^2}\right)^{-(t_c/\varepsilon+\tau_\varepsilon)/2} = \left(1 + \frac{t^2}{t_c^2}\right)^{-\tau_\varepsilon/2} \left(e^{-\frac{t^2}{2\varepsilon t_c}} + O\left(\varepsilon e^{-\frac{t^2}{4\varepsilon t_c}}\right)\right),$$

where the last equality holds from using lemma 3 with $x = t^2/t_c^2 \in [0, 1]$ and $k = t_c/2\varepsilon$. This term bounds the modulus of $\mathrm{Re}(f^{n_\varepsilon+1})$, and bounds $R^\alpha_{n_\varepsilon+1}$ up to a factor of $(n_\varepsilon+1)^{-\alpha}$ again, which is $O(1)$. So we have bounded both term on the left hand side of (50). For $|t| \ge \varepsilon^{\alpha/2}$, note that $e^{-t^2/2\varepsilon t_c} = O\left(\varepsilon e^{-t^2/4\varepsilon t_c}\right) = O\left(\varepsilon^\alpha e^{-t^2/4\varepsilon t_c}\right)$, so both terms go in the error and we are done.

For $|t| < \varepsilon^{\alpha/2}$, the prefactor $(1 + t^2/t_c^2)^{-\tau_\varepsilon/2} = 1 + O(\varepsilon^\alpha)$, thus giving

$$|f|^{n_\varepsilon+1} = e^{-\frac{t^2}{2\varepsilon t_c}} + O\left(\varepsilon^\alpha e^{-\frac{t^2}{4\varepsilon t_c}}\right).$$

In exactly the same way,

$$R^\alpha_{n_\varepsilon+1} = (n_\varepsilon+1)^{-\alpha}|f|^{n_\varepsilon+1} = O(\varepsilon^\alpha)\left(e^{-\frac{t^2}{2\varepsilon t_c}} + O\left(\varepsilon^\alpha e^{-\frac{t^2}{4\varepsilon t_c}}\right)\right),$$

which goes in the error ϕ^α, observing that $e^{-\frac{t^2}{2\varepsilon t_c}} = O(e^{-\frac{t^2}{4\varepsilon t_c}})$ for all values of t. We now require analysis of the phase of $f^{n_\varepsilon+1}$, for which it is convenient to write f as the sum of its real and imaginary part:

$$f(t) = \frac{t_c^2}{t^2+t_c^2} + i\frac{tt_c}{t^2+t_c^2}.$$

This gives that the relevant phase is

$$\begin{aligned}\exp\{i(n_\varepsilon+1)\arctan(t/t_c)\} &= \exp\left\{i\left(\frac{t_c}{\varepsilon}+\tau_\varepsilon\right)\left(t/t_c - \frac{1}{3}(t/t_c)^3 + O\left((t/t_c)^5\right)\right)\right\},\\ &= \exp\left\{i\left(\frac{t}{\varepsilon} - \frac{t^3}{3\varepsilon t_c^2} + \frac{\tau_\varepsilon t}{t_c}\right) + O\left(\tau_\varepsilon (t/t_c)^3\right) + O\left((t/t_c)^5/\varepsilon\right)\right\},\end{aligned}$$

where we used the Taylor expansion of arctan. We can now use $e^x = 1 + O(x)$ and $|t| \le \varepsilon^{\alpha/2}$ to obtain

$$\exp\left\{O\left(\tau_\varepsilon (t/t_c)^3\right) + O\left((t/t_c)^5/\varepsilon\right)\right\} = 1 + O(\varepsilon^{3\alpha/2}) + O(\varepsilon^{5\alpha/2-1}),$$

thus yielding

$$\mathrm{Re}(f^{n_\varepsilon+1}) = \left(e^{-\frac{t^2}{2\varepsilon t_c}} + O\left(\varepsilon^\alpha e^{-\frac{t^2}{4\varepsilon t_c}}\right)\right)\cos\left(\frac{t}{\varepsilon} - \frac{t^3}{3\varepsilon t_c^2} + \frac{\tau_\varepsilon t}{t_c}\right)\left(1 + O(\varepsilon^{3\alpha/2}) + O(\varepsilon^{5\alpha/2-1})\right),$$

ending the proof of the theorem. □

6.2 Unitary evolution operator

We now use standard perturbation theory to calculate the unitary evolution operator in the optimal superadiabatic basis; the next proposition is exactly Corollary 1 in [BT05a] and follows the discussion around Theorem X.69 in [RS75].

Recall the error function $\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-x^2}\mathrm{d}x$, which smoothly and monotonically increases from -1 to 1. Similarly to [BT05a], we define $\Delta(t,s) = \theta(t) - \theta(s)$.

Proposition 3. *Let $K_\varepsilon^{n_\varepsilon}(t,s)$ be the unitary evolution operator in the superadiabatic basis,*

that is the solution to

$$i\varepsilon\frac{\mathrm{d}}{\mathrm{d}t}K_\varepsilon^{n_\varepsilon}(t,s) = H_\varepsilon^{n_\varepsilon}K_\varepsilon^{n_\varepsilon}(t,s) \text{ such that } K_\varepsilon^{n_\varepsilon}(t,t) = I_2. \tag{51}$$

Then

$$K_\varepsilon^{n_\varepsilon}(t,s) = \begin{pmatrix} k_\varepsilon^+(t,s) & k_\varepsilon(t,s) \\ \overline{k}_\varepsilon(t,s) & k_\varepsilon^-(t,s) \end{pmatrix},$$

where

$$k_\varepsilon^\pm(t,s) = e^{\mp\frac{i(t-s)}{2\varepsilon}} + O(\varepsilon\Delta(t,s)) \tag{52}$$

and

$$k_\varepsilon(t,s) = \sin(\frac{\pi\gamma}{2})e^{-\frac{t_c}{\varepsilon}}e^{-\frac{i(t+s)}{2\varepsilon}}\left(\operatorname{erf}\left(\frac{t}{\sqrt{2\varepsilon t_c}}\right) - \operatorname{erf}\left(\frac{s}{\sqrt{2\varepsilon t_c}}\right)\right) + O(\varepsilon^\alpha e^{-t_c/\varepsilon}\Delta(t,s)) \tag{53}$$

for $|t|, |s| > \varepsilon^{\beta/2}$ *for some* $\beta \in (0,1)$.

Proof. The proof goes as follows, see [RS75] for a more expansive discussion applied to a more general context. We can split the Hamiltonian into

$$H_\varepsilon^{n_\varepsilon} = H_0 + V_\varepsilon^{n_\varepsilon}, \text{ where } H_0 = \frac{1}{2}\sigma_z, \text{ and } V_\varepsilon^{n_\varepsilon} = \begin{pmatrix} O(\varepsilon^2\dot{\theta}) & c_\varepsilon^{n_\varepsilon} \\ \overline{c}_\varepsilon^{n_\varepsilon} & O(\varepsilon^2\dot{\theta}) \end{pmatrix}.$$

Now define $\tilde{V}_\varepsilon^{n_\varepsilon}(t) = e^{itH_0/\varepsilon}V_\varepsilon^{n_\varepsilon}(t)e^{-itH_0/\varepsilon}$ and let $\tilde{K}_\varepsilon^{n_\varepsilon}(t,s)$ solve $i\varepsilon\frac{\mathrm{d}}{\mathrm{d}t}\tilde{K}_\varepsilon^{n_\varepsilon}(t,s) = \tilde{V}_\varepsilon^{n_\varepsilon}(t)\tilde{K}_\varepsilon^{n_\varepsilon}(t,s)$. Letting $K_\varepsilon^{n_\varepsilon}(t,s) = e^{-itH_0/\varepsilon}\tilde{K}_\varepsilon^{n_\varepsilon}(t,s)e^{isH_0/\varepsilon}$, one can easily check that it solves (51), and we will follow this construction.

Obviously,

$$\tilde{K}_\varepsilon^{n_\varepsilon}(t,s) = \exp\left(\frac{-i}{\varepsilon}\int_s^t \tilde{V}_\varepsilon^{n_\varepsilon}(u)\mathrm{d}u\right),$$

so we will require analysis of this exponential, and we focus on its argument. We can easily compute

$$\tilde{V}_\varepsilon^{n_\varepsilon} = \begin{pmatrix} O(\varepsilon^2\dot{\theta}) & e^{it/\varepsilon}c_\varepsilon^{n_\varepsilon} \\ e^{-it/\varepsilon}\overline{c}_\varepsilon^{n_\varepsilon} & O(\varepsilon^2\dot{\theta}) \end{pmatrix},$$

thus the diagonal elements of $-i/\varepsilon\int\tilde{V}_\varepsilon^{n_\varepsilon}$ are $O(\varepsilon\Delta(t,s))$. The main challenge is to calculate the off-diagonal elements

$$\begin{aligned} -\frac{i}{\varepsilon}\int_s^t e^{\frac{iu}{\varepsilon}}c_\varepsilon^{n_\varepsilon}(u)\mathrm{d}u = \sqrt{\frac{2}{\varepsilon\pi t_c}}\sin(\gamma\pi/2)e^{-t_c/\varepsilon}\int_s^t e^{-\frac{u^2}{2\varepsilon t_c}}\left(e^{\frac{2iu}{\varepsilon}-\frac{iu^3}{3\varepsilon t_c^2}+\frac{i\tau_\varepsilon u}{t_c}} + e^{\frac{iu^3}{3\varepsilon t_c^2}-\frac{i\tau_\varepsilon u}{t_c}}\right)\mathrm{d}u \\ + O\left(\varepsilon^{-1/2}e^{-t_c/\varepsilon}\int_s^t e^{\frac{iu}{\varepsilon}}\phi^\alpha(\varepsilon,u)\mathrm{d}u\right), \end{aligned}$$

where we used (46) from Theorem 3, $\cos x = (e^x + e^{-x})/2$ and multiplied out the $e^{iu/\varepsilon}$. Now using $e^x - 1 - x = O(x^2)$ for $x = \pm\left(\frac{iu^3}{3\varepsilon t_c^2} - \frac{i\tau_\varepsilon u}{t_c}\right)$, we replace e^x by $1+x$ and the first

integral becomes

$$\int_s^t e^{-\frac{u^2}{2\varepsilon t_c}} \left(e^{\frac{2iu}{\varepsilon}} \left(1 - \frac{iu^3}{3\varepsilon t_c^2} + \frac{i\tau_\varepsilon u}{t_c} \right) + \left(1 + \frac{iu^3}{3\varepsilon t_c^2} - \frac{i\tau_\varepsilon u}{t_c} \right) \right) \mathrm{d}u + \text{error},$$

where the error is bounded by a constant times the integral

$$\int_{-\infty}^{\infty} e^{-\frac{u^2}{2\varepsilon t_c}} (u^6/\varepsilon^2 + u^4/\varepsilon + u^2) \mathrm{d}u = O(\varepsilon^{3/2}).$$

We have

$$\int_s^t e^{-\frac{u^2}{2\varepsilon t_c}} \mathrm{d}u = \sqrt{\frac{\varepsilon \pi t_c}{2}} \left(\operatorname{erf}\left(\frac{t}{\sqrt{2\varepsilon t_c}} \right) - \operatorname{erf}\left(\frac{s}{\sqrt{2\varepsilon t_c}} \right) \right),$$

and all other terms integrate to order $O(\varepsilon^{\alpha+1/2}\Delta(t,s))$ if $|t|, |s| > \varepsilon^{\beta/2}$ for some $\beta \in (0,1)$. Similar calculations for the integral involving ϕ^α show it is also of order $O(\varepsilon^{\alpha+1/2}\Delta(t,s))$ for the relevant values of t, s. The off-diagonal elements of $-i/\varepsilon \int \tilde{V}_\varepsilon^{n_\varepsilon}$ are therefore

$$\sin(\gamma\pi/2) e^{-t_c/\varepsilon} \left(\operatorname{erf}\left(\frac{t}{\sqrt{2\varepsilon t_c}} \right) - \operatorname{erf}\left(\frac{s}{\sqrt{2\varepsilon t_c}} \right) \right) + O(\varepsilon^\alpha e^{-t_c/\varepsilon} \Delta(t,s)) = O(\varepsilon^\alpha e^{-t_c/\varepsilon} \Delta(t,s)). \tag{54}$$

We now calculate the relevant matrix exponential, for which we want to consider the matrix power series and bound the terms of order larger than or equal to 2. To this end, considering the off-diagonal elements as in the right hand side of (54), one can inductively check that for $n \geq 2$,

$$\left(\frac{-i}{\varepsilon} \int_s^t \tilde{V}_\varepsilon^{n_\varepsilon} \right)^n = (\Delta(t,s))^n \begin{pmatrix} O(\varepsilon^2) & O(\varepsilon^\alpha e^{-t_c/\varepsilon}) \\ O(\varepsilon^\alpha e^{-t_c/\varepsilon}) & O(\varepsilon^2) \end{pmatrix},$$

and noting that $\sum_{n=2}^\infty x^n/n! \leq x^2 e^x$ for $x \geq 0$ and that $\Delta(t,s)$ is bounded uniformly in t, s, we have that $\sum_{n=2}^\infty (\Delta(t,s))^n/n! = O(\Delta(t,s))$. Turning back to the matrix exponential, this yields

$$\exp\left(\frac{-i}{\varepsilon} \int_s^t \tilde{V}_\varepsilon^{n_\varepsilon} \right) = I_2 - \frac{i}{\varepsilon} \int_s^t \tilde{V}_\varepsilon^{n_\varepsilon} + \Delta(t,s) \begin{pmatrix} O(\varepsilon^2) & O(\varepsilon^\alpha e^{-t_c/\varepsilon}) \\ O(\varepsilon^\alpha e^{-t_c/\varepsilon}) & O(\varepsilon^2) \end{pmatrix}.$$

We then use the left hand side of (54), multiply from the left by $e^{-itH_0/\varepsilon}$ and from the right by $e^{isH_0/\varepsilon}$ to get the result. □

One can replace ε^α by $\sqrt{\varepsilon}$ in (53) to get a bound that holds uniformly in t, s. This then implies the existence of solutions to the Schrödinger equation of the form

$$U_\varepsilon^{n_\varepsilon}(t) \begin{pmatrix} e^{-it/2\varepsilon} \left(1 + O(\varepsilon(\arctan(t) + \pi/2))\right) \\ \sin(\gamma\pi/2) e^{-t_c/\varepsilon} e^{it/2\varepsilon} \left(\operatorname{erf}\left(\frac{t}{\sqrt{2\varepsilon t_c}} \right) + 1 \right) \end{pmatrix} + O(\sqrt{\varepsilon} e^{-t_c/\varepsilon}),$$

which start at the ground state for large negative times, and develop an exponentially small component of the excited state with an error function. In [Ber90] and [BL93], Berry and Lim argue that this is universal, that is solutions like the above exist without depending on the detailed time dependence of the Hamiltonian. Recall that the only choice we made for the Hamiltonian was the form of $\dot{\theta}$, and one can show that we have

$$\dot{\theta}(t) = \frac{\pm i\gamma}{t \pm it_c} + O(|t \pm it_c|^a)$$

for some $a < -1$ close to its singularities nearest to the real axis, in a broad genericity class, as mentioned in [BT05b].

6.3 The scattering map

We now have all the pieces to define and calculate the scattering map. As mentioned earlier, the scattering map takes solutions to the unperturbed, autonomous equation at $-\infty$ and sends them to the corresponding solution at $+\infty$. More precisely, let $\psi_\pm(t)$ be solutions to $i\varepsilon\dot{\psi}_\pm(t) = H_0\psi_\pm(t)$, where $H_0 = \frac{1}{2}\sigma_z$, and one should think of ψ_- as the behaviour of a solution to (4) at time $-\infty$ and of ψ_+ as that at time $+\infty$. For fixed ε, we want to consider the map S_ε such that $S_\varepsilon(\psi_-(0)) = \psi_+(0)$, and we go through the heuristics behind our definition. From the definition of ψ_+, so we have $\psi_+(0) = e^{iH_0t/\varepsilon}\psi_+(t)$, and now comes the crucial point: for $|t|$ large, the perturbed problem essentially behaves like the unperturbed problem, so if we let K denote the unitary evolution operator for the perturbed equation and let K_0 be that of the unperturbed equation, we have $K(t,-t) \approx K_0(t,-t)$. We then write $e^{iH_0t/\varepsilon}\psi_+(t) \approx e^{iH_0t/\varepsilon}K(t,-t)\psi_-(-t)$, and we use the definition of ψ_- to get $S(\psi_-(0)) \approx e^{iH_0t/\varepsilon}K(t,-t)e^{iH_0t/\varepsilon}\psi_-(0)$, where this is a "good" approximation for $|t|$ large. We define the scattering matrix to be

$$S_\varepsilon := \lim_{t\to\infty} e^{iH_0t/\varepsilon}K(t,-t)e^{iH_0t/\varepsilon}, \tag{55}$$

and computing this is now a simple corollary of Proposition 3.

Corollary 1. *For $\alpha \in (0,1)$, we have*

$$S_\varepsilon = \begin{pmatrix} 1+O(\varepsilon) & 2\sin(\frac{\gamma\pi}{2})e^{-t_c/\varepsilon}(1+O(\varepsilon^\alpha)) \\ 2\sin(\frac{\gamma\pi}{2})e^{-t_c/\varepsilon}(1+O(\varepsilon^\alpha)) & 1+O(\varepsilon) \end{pmatrix}.$$

Proof. We will change to the optimal superadiabatic basis using $U_\varepsilon^{n_\varepsilon}$, and recall from Lemma 2 that we have $\|U_\varepsilon^{n_\varepsilon}(t) - I_2\| \to 0$ as $|t| \to \infty$. Thus

$$S_\varepsilon = \lim_{t\to\infty} e^{iH_0t/\varepsilon}U_\varepsilon^{n_\varepsilon *}(t)K(t,-t)U_\varepsilon^{n_\varepsilon}(-t)e^{iH_0t/\varepsilon} = \lim_{t\to\infty} e^{iH_0t/\varepsilon}K_\varepsilon^{n_\varepsilon}(t,-t)e^{iH_0t/\varepsilon},$$

where the result follows from Proposition 3. □

The transition amplitude $A(\varepsilon)$ for going from the ground state to the excited state[18] is thus given by looking at the $(1,2)$-th entry of S_ε:

$$A(\varepsilon) = 2\sin(\frac{\gamma\pi}{2})e^{-t_c/\varepsilon}(1 + O(\varepsilon^\alpha))$$

for every $\alpha \in (0,1)$, exactly as in [BT05a]. Squaring gives the transition probability, which one can check agrees with the results in [BL93] and [Joy93].

7 Conclusion

We explored 2-level quantum systems with slowly changing Hamiltonians, for which the Schrödinger equation and von Neumann equations reduce to (4) and (5) respectively, with Hamiltonian given by (7) after a few simplifications. This is simplified further by using the adiabatic basis – consisting of eigenvectors of H – and the Hamiltonian becomes almost diagonal, cf (8). However, this also presented us with a choice of coupling function $\dot{\theta}$ in (9), so the case studied is not entirely general.

Focussing on the von Neumann equation, we use the well known Pauli spin matrices to present a completely equivalent equation (11) on the unit 2-sphere, where the dynamics look like fast rotations around a changing axis of rotation. The next step required a discussion on perturbation theory, which justified writing solutions as power series in the adiabatic parameter (12), then a small bit of algebra gave us a scalar recursion for the components of the projections.

An insight given by complex analysis helped study this recursion; it gave us an explicit formula for each component in Proposition 1 and a statement concerning the asymptotics of y_{n+1} in Theorem 1. Studying the behaviour of the coefficients in the asymptotic series informed us of its divergence, and motivated considering not the full series, but rather a truncated sum. Going back to projections, the truncated sum $\pi^{(n)}$ gives an approximate solution to the von Neumann equation which is an error of order ε^{n+1} away from being a solution. From here we construct a special unitary matrix consisting of eigenvectors of $\pi^{(n)}$, which sends us to a so-called superadiabatic basis, where the Hamiltonian has off-diagonal elements of order ε^{n+1}. The notion of the optimal index at which truncate was then crucial to obtain the smallest off-diagonal elements, and choosing an index near t_c/ε gives that they are exponentially small.

Another use of perturbation theory then yields the unitary evolution operator to the Schrödinger equation, immediately giving the scattering map, as well as the transition amplitude of going from the ground state to the excited state.

Betz and Teufel's work [BT05a] and this dissertation rigorously obtains the transition probability described in [Ber90] and [Joy93], for a 2-dimensional Hamiltonian with coupling

[18] Recall that the ground state is $v_2 = \binom{0}{1}$ and the excited state is $v_1 = \binom{1}{0}$.

θ in a restricted form. In particular, the main objective was to calculate the scattering, for which we calculated the evolution operator, essentially giving the complete behaviour of solutions to the Schrödinger equation.

There are a few improvements that could follow from this research project. Firstly, one could present and prove a statement about the scattering exclusively using the sphere equation, without going back to the Schrödinger equation and/or without using eigenvectors of projections to almost diagonalise the Hamiltonian, but rather finding an equivalent procedure for (11). Note that an almost diagonal H for the von Neumann is equivalent to an almost vertical $\mathbf{a}$ for the sphere equation. Second, one could treat the equation obtained on the plane by stereographic projection, as discussed in Appendix A. The main advantage of the plane equation is that is it just a scalar equation, with the price for that being the loss of linearity. One can then guess the scattering map would take the form of an $SU(2)$ Möbius transformation. Lastly, one could modify this proof and consider generic forms for the coupling $\dot{\theta}$, to rigorously prove the result of [Ber90] for a broad genericity class of Hamiltonians.

A On the complex plane

With an equation on the 2-dimensional unit sphere, a natural next step is to use stereographic projection to get one on the plane. More precisely, define $w = \frac{x+iy}{1-z}$. Geometrically, this is the projection onto the plane $\{z = -1\}$, with the North pole at $(0, 0, 1)$. Writing

$$\pi = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} = \frac{1}{2}(\sigma_0 + x\sigma_x + y\sigma_y + z\sigma_z),$$

one gets

$$w = \frac{\pi_{11}}{\pi_{12}} = \frac{\pi_{21}}{\pi_{22}}, \tag{56}$$

and a bit of algebra shows that w solves

$$i\varepsilon\dot{w} = \varepsilon\dot{\theta}w^2 + w - \varepsilon\dot{\theta}. \tag{57}$$

Observe that $-1/\overline{w}$ corresponds to the antipodal point on the sphere, and satisfies (57) as well. Using this gives us the behaviour in a neighbourhood around infinity, so we understand the whole Riemann sphere.

One obvious advantage is that this is a scalar equation, whereas all the others are at least vector valued. This does come at the cost of losing linearity with the quadratic term.

We can also write π as a function of w:

$$\pi(w) = \frac{1}{1 + w\overline{w}} \begin{pmatrix} w\overline{w} & \overline{w} \\ w & 1 \end{pmatrix}, \tag{58}$$

and calculations also show that the unitary matrix

$$V(w) = \frac{1}{\sqrt{1 + w\overline{w}}} \begin{pmatrix} \overline{w} & -1 \\ 1 & w \end{pmatrix} \tag{59}$$

diagonalises π to $P_1 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, that is $\pi V = VP_1$.

Another way of obtaining w is through the Schrödinger equation (4). For a solution $\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$, one can check that $w = \psi_1/\psi_2$ solves (57) as well. Observe that multiplying ψ by any arbitrary phase factor $e^{i\theta}$ does not change w, just as when considering the projection π_ψ onto ψ. So while equations (4) and (5) are not equivalent because of this possible phase, (5) and (57) are. Note that $\frac{-1}{\overline{w}}$ is the ratio of the components of $\psi^\perp = \begin{pmatrix} -\overline{\psi}_2 \\ \overline{\psi}_1 \end{pmatrix}$, reinforcing the notion that antipodal points on the sphere correspond to orthogonal subspaces.

Let $K = \exp(-i\int_s^t H) \in SU(2)$ be the evolution operator for the Schrödinger equation, and let us write it as $K = \begin{pmatrix} \lambda & -\overline{\mu} \\ \mu & \overline{\lambda} \end{pmatrix}$. Then any solution to (57) is of the form $\frac{\lambda w_0 - \overline{\mu}}{\mu w_0 + \overline{\lambda}}$ for some constant $w_0 \in \mathbb{C}$, that is the evolution operator acts on solutions through Möbius transformations on the Riemann sphere. Observe that we can map any element of $SU(2)$

to an element of $SO(3)$: writing $\lambda = a + bi$ and $\mu = c + di$, take

$$U \mapsto \begin{pmatrix} 1 - 2b^2 - 2c^2 & 2cd - 2ab & 2ac + 2bd \\ 2ac + 2cd & 1 - 2b^2 - 2d^2 & 2bc - 2ad \\ 2bd - 2ac & 2ad + 2bc & 1 - c^2 - d^2 \end{pmatrix}.$$

This group homomorphism is 2-to-1 as U and $-U$ map to the same orthogonal matrix, but considering the induced Möbius transformation gives a bijective relation. This then gives the evolution operator R for the equation (11) on the 2-sphere.

We now look at the corresponding asymptotic series. Note that there are no conditions on w, so we will only require the recursion, another bonus of this equation. For

$$w = \sum_{k \geq 0} \varepsilon^k w_k,$$

we have

$$\begin{aligned} w_0 &= 0, \\ i\dot{w}_0 &= \dot{\theta} w_0^2 + w_1 - \dot{\theta}, \\ i\dot{w}_k &= \dot{\theta} \sum_{j=0}^{k} w_j w_{k-j} + w_{k+1}, \text{ for } k \geq 1. \end{aligned}$$

One of the main advantages of this recursion is that we have no condition, which comes from the fact that the space of interest is linear, so any linear combination will stay in it. Indeed, any truncated sum $w^{(n)} = \sum_{k=0}^{n} \varepsilon^k w_k$ will obviously be in $\mathbb{C}$, in contrast to (11), where truncating the series only means we are an error of order ε^{n+1} away from being on the sphere.

One could observe that in the other recursions, we had 2 possible starting points $\mathbf{x}_0$ or π_0, but there is only one here. In fact, we need to consider the equation for the antipodal point $-1/\overline{w}$ to get the full picture, which gives $w_0 = \infty$, so we do have 2 starting points, each corresponding to an eigenspace of H_0.

Bibliography

References

[AS64] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, ninth dover printing, tenth gpo printing edition, 1964.

[BB14] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical Computer Science*, 560:7–11, 2014. Theoretical Aspects of Quantum Cryptography – celebrating 30 years of BB84.

[Ber90] Michael V. Berry. Histories of adiabatic quantum transitions. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 429(1876):61–72, 1990.

[BF28] Max Born and V. Fock. Beweis des Adiabatensatzes. (German) [Proof of the Adiabatic Theorem]. 51(3–4):165–180, March 1928.

[BL93] M. V. Berry and R. Lim. Universal transition prefactors derived by superadiabatic renormalization. *J. Phys. A 26*, pages 4737–4747, 1993.

[BT05a] Volker Betz and Stefan Teufel. Precise coupling terms in adiabatic quantum evolution. *Annales Henri Poincaré*, 6(2):217–246, apr 2005.

[BT05b] Volker Betz and Stefan Teufel. Precise coupling terms in adiabatic quantum evolution: The generic case. *Communications in Mathematical Physics*, 260(2):481–509, aug 2005.

[CJ04] Dariusz Chruściński and Andrzej Jamiołkowski. Geometric phases in classical and quantum mechanics. 2004.

[Din73] Robert B. Dingle. *Asymptotic Expansions: Their Derivation and Interpretation*. Academic Press, 1973.

[Gel18] Vassili Gelfreich. Lecture notes for MA4A7 Quantum Mechanics: Basic Principles and Probabilistic Methods, 2018.

[HJ04] George A. Hagedorn and Alain Joye. Time development of exponentially small non-adiabatic transitions. *Communications in Mathematical Physics*, 250(2):393–413, jul 2004.

[HJ05] George A. Hagedorn and Alain Joye. Recent results on non–adiabatic transitions in quantum mechanics, 2005.

[Joy93] Alain Joye. Non-trivial prefactors in adiabatic transition probabilities induced by high order complex degeneracies. *J. Phys. A 26*, pages 6517–6540, 1993.

[Kat50] Tosio Kato. On the adiabatic theorem of quantum mechanics. *Journal of the Physical Society of Japan*, 5(6):435–439, 1950.

[MM04] Dan Marinescu and Gabriela Marinescu. *Approaching Quantum Computing*. 2004.

[RS75] Micheal Reed and Barry Simon. *Method of Modern Mathematical Physics II: Fourier Analysis, Self-Adjointness*. Academic Press, 1975.

[RX11] Juanjo Rué and Sebastian Xambó. Mathematical essentials of quantum computing. 2011.

[Sho94] P.W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, 1994.

[Teu03] Stefan Teufel. *Adiabatic perturbation theory in quantum dynamics*. Springer Science & Business Media, 2003.