Geometry of the adiabatic theorem

Augusto César Lobo¹, Rafael Antunes Ribeiro², Clyffe de Assis Ribeiro³ and Pedro Ruas Dieguez³

¹ Departamento de Física, Instituto de Ciências Exatas e Biológicas, Universidade Federal de Ouro Preto, CEP 35400-000, Ouro Preto, Minas Gerais, Brazil
² Centro Federal de Educação Tecnológica de Minas Gerais Campus VIII, CEP 37022-560, Varginha, Minas Gerais, Brazil
³ Departamento de Física, Instituto de Ciências Exatas, Universidade Federal de Minas Gerais, CP 702, CEP 30161-970, Belo Horizonte, Minas Gerais, Brazil

E-mail: lobo@iceb.ufop.br

Received 9 March 2012, in final form 4 May 2012
Published 7 June 2012
Online at stacks.iop.org/EJP/33/1063

Abstract

We present a simple and pedagogical derivation of the quantum adiabatic theorem for two-level systems (a single qubit) based on geometrical structures of quantum mechanics developed by Anandan and Aharonov, among others. We have chosen to use only the minimum geometric structure needed for the understanding of the adiabatic theorem for this case.

(Some figures may appear in colour only in the online journal)

1. Introduction

The quantum mechanical adiabatic theorem is one of the most important and fruitful tools of quantum physics. It was first stated in 1928 by Born and Fock [1] and set in a more rigorous mathematical foundation in 1950 by Kato [2]. In the mid-1980s, it was linked in a new and fundamental way with a more geometrical view of quantum mechanics by the so-called Berry’s phase or geometric phase [3, 4].

More recently, it has received a renewal of interest because of the advent of quantum information theory and in particular, the quantum adiabatic computation model [5, 6]. In fact, some authors have questioned the theorem and have tried to look for a more precise statement and conditions for its consistency and validity [7–15].

Though the theorem can be stated and understood quite easily and it is, in fact, routinely discussed in undergraduate textbooks [16, 17], we believe that the reason why it works, though, is less known. It can be simply stated as saying that under a certain special context, if the state of a quantum mechanical system is an eigenstate of its Hamiltonian, it will remain as such. The special context means that the time-dependent Hamiltonian must change slowly in some sense (adiabaticity) if the spectrum obeys some technical conditions. Stated in this manner, one can say it is a reasonably intuitive assertion and can be traced back to its corresponding classical theorem on adiabatic invariants [18]. However, the proofs available to students such
section 3, we conclude with some closing remarks. In a very clean and pedagogical geometric setting which we believe improves the physical understanding of this important theorem for students of physics and even for specialists. In section 3, we conclude with some closing remarks.

2. The geometry of quantum evolution

Let $W^{n+1}$ be an $(n + 1)$-dimensional Hilbert space together with its dual $W^{n+1}$ and $\{ |u_0\} (\sigma = 0, 1, \ldots, n)$ also be an arbitrary basis for $W^{n+1}$. A Hermitian inner product may be introduced by anti-linear mapping $\dagger : W^{n+1} \rightarrow W^{n+1}$ (where $\dagger$ is the familiar ‘dagger’ operation). Indeed, the inner product between two arbitrary states $|\psi\rangle$ and $|\phi\rangle$ can now be defined as

\[ \langle \psi | \phi \rangle = |\psi\rangle \langle \phi | \dagger \rangle = \langle \psi | \phi \rangle. \]

Thus, an arbitrary normalized ket $|\psi\rangle$ expanded in such a basis can be represented by a complex $(n + 1)$-column matrix:

\[ |\psi\rangle = \sum |u_\sigma\rangle \psi^\sigma = (\psi_0 \ 0 \ \psi_1 \ 0 \ \ldots \ \psi_n)^T, \quad \text{with} \quad \sum \overline{\psi}_\sigma \psi^\sigma = 1. \quad (1) \]

Writing the complex amplitudes as $\psi^\sigma = x^\sigma + iy^\sigma$, one can easily see that the set of normalized states can be identified with a $(2n + 1)$-dimensional sphere $S^{2n+1}$ contained in $W^{n+1}$. Since two state vectors that differ by a complex phase cannot be physically distinguished by any means, it is convenient to define the true physical space of states as the above set of normalized states modulo the equivalence relation in $S^{2n+1}$ defined in the following way. We say that $|\psi\rangle$ is equivalent to $|\phi\rangle$ if and only if there is a real number $\theta$ such that $|\psi\rangle = e^{i\theta} |\phi\rangle$.

The space of rays defined above is also known as the $n$-dimensional (complex) projective space $\mathbb{CP}(n)$. A standard complex coordinate system for $\mathbb{CP}(n)$ is provided by $n$ complex numbers $\xi^i = \psi^i/\psi^0$ ($i = 1, \ldots, n$) for those points where $\psi^0 \neq 0$. In the $n = 1$ case, we have a single qubit described by a single complex coordinate $\xi$. In this case, $\mathbb{CP}(1)$ is topologically equivalent to a two-dimensional sphere and the stereographic projection map shows explicitly

\[ \xi = \tan(\theta/2) e^{i\phi}. \quad (2) \]

(see figure 1) that $\mathbb{CP}(1)$ can be seen as the complex plane ‘plus’ a point at ‘infinity’.

This description provides us with the so-called Bloch sphere (or Riemann sphere) with standard coordinates. Thus, any physical state can be expressed as a normalized state represented as a point on the Bloch sphere in the following standard form:

\[ |\psi\rangle = |\theta, \phi\rangle = \cos(\theta/2) |u_0\rangle + e^{i\phi} \sin(\theta/2) |u_1\rangle, \quad (3) \]

where one can easily see that antipodal points in the Bloch sphere represent orthogonal state vectors. A natural metric $\mathbb{CP}(n)$ can be defined in the following way. Let points $P_0$ and $P_1 \in \mathbb{CP}(n)$ be projections respectively from two infinitesimally nearby normalized state vectors $|\psi\rangle$ and $|\psi + d\psi\rangle$. It is then natural to define the squared distance between $P_0$ and $P_1$ as the projection of $|d\psi\rangle$ in the ‘orthogonal direction’ of $|\psi\rangle$, that is, the projection given by the projection operator $\hat{P}_\perp = \hat{I} - |\psi\rangle\langle\psi|$ as shown in figure 2.
It is then easy to see that (by the idempotence of $\hat{\pi}_\perp$)
\[
ds^2(CP(n)) = |(d\psi|\hat{\pi}_\perp^\dagger d\psi)|^2 = (d\psi|d\psi) - (d\psi|\psi)(\psi|d\psi).
\] (4)

Let $|\psi(t)\rangle$ be the curve of normalized state vectors in $W^{n+1}$ given by the unitary evolution generated by a Hamiltonian $\hat{H}$. The Schrödinger equation implies a relation between $|\psi(t)\rangle$ and $|\psi(t + dt)\rangle$, given by
\[
|d\psi| = |\psi(t + dt)| - |\psi(t)| = -i\hat{H}|\psi(t)\rangle dr.
\] (5)

For the sake of simplicity, we have used and henceforth will be assuming $\hbar = 1$ units. The above equation together with (4) gives us the squared distance between two infinitesimally nearby projections of state vectors connected by the unitary evolution over $CP(n)$:
\[
ds^2(CP(n)) = \left(\langle \psi(t)|\hat{H}|\psi(t)\rangle - \langle \psi(t)|\hat{H}|\psi(t)\rangle \right)^2 dr^2 = g^2_{\psi(t)} E dr^2.
\] (6)
In particular, for a single qubit, one can write the metric over the $\mathbb{CP}(1)$ sphere by (3) as
\[
d s^2 = \frac{1}{4} (d\theta^2 + \sin^2 \theta d\varphi^2),
\]
where we can immediately identify the Bloch sphere as a two-dimensional sphere with radius $r = 1/2$. Note that some authors define the Bloch sphere as a unit radius sphere. Our choice seems more natural to us because of the above geometrical construction and also in order to identify the Bloch sphere with the so-called Riemann sphere.

2.1. The time–energy uncertainty relation

Equation (6) leads to a very elegant relation between the speed of the projection over $\mathbb{CP}(n)$ and the instantaneous energy uncertainty [19]:
\[
\frac{ds}{dt} = \delta E(t).
\]
The above equation contains in itself the seeds of both a beautiful geometric formulation of the time–energy uncertainty relation and the adiabatic theorem. Indeed, for the latter, it is quite clear that if a state is initially (let us say for $t = 0$) an energy eigenstate, then one has $\delta E(0) = 0$ and the instantaneous speed must then also be null. If one moves the state very slowly around $\mathbb{CP}(n)$, such that $ds/dt \approx 0$, then one should have in a self-consistent way that $\delta E(t) \approx 0$ for all $t$. But this means that the state must remain an eigenstate for all $t$, which is exactly a statement of the adiabatic theorem and we shall discuss this issue further in the following subsection.

Before that, let us close this subsection by observing how equation (8) also implies a geometric version of the time–energy uncertainty relation due to Anandan and Aharonov. Let us assume at first that the system consists only of a single qubit. Let $P_0$ and $P_1$ be two distinct points on Bloch sphere located on an arbitrary path driven by a time-dependent Hamiltonian $\hat{H}(t)$ at instants respectively $t_0$ and $t_1$. One can define the time average of the energy uncertainty $\overline{\delta E}$ between $t_0$ and $t_1$ as
\[
\overline{\delta E} = \frac{1}{\Delta t} \int_{t_0}^{t_1} \delta E(t) \, dt \quad \text{with} \quad \Delta t = t_1 - t_0.
\]
With this definition and from (8), it is easy to see that
\[
\Delta s = \overline{\delta E} \Delta t.
\]
Now to derive the time–energy uncertainty relation, we may borrow an elementary argument from quantum mechanics updated with a more modern flavour from quantum information. In fact, given a particular state (an arbitrary point in the Bloch sphere), one can physically distinguish it from another state if and only if this second point is orthogonal to it (its antipodal point on the sphere). In fact, the indistinguishability of non-orthogonal state vectors is a basic result in quantum information together with the no-cloning theorem, for instance. The indistinguishability of non-orthogonal states means that one party (let us say Alice) cannot reliably transmit information to the second party (call him Bob) from a common agreed upon alphabet of non-orthogonal states because these states cannot be simultaneously eigenstates of commuting observables. For two-level systems, this means that for the two points $P_0$ and $P_1$ to be physically distinguished, their minimum distance must necessarily be the half-length of a great circle of the Bloch sphere, which is clearly $\pi/2$. So we may write the time–energy uncertainty relation as
\[
\overline{\delta E} \Delta t \geq \pi/2.
\]
Note that the above inequality differs from the usual textbook relation by a factor of \( \pi \). This is a very different mathematical picture than the one exhibited by the celebrated Heisenberg uncertainty relation for position and momentum. But this is expected because in non-relativistic mechanics, time is an external parameter and not an observable as it happens to be the case for the position operator. In fact, some authors refer to the time \( \Delta t \) in the inequality above as passage time [20].

2.2. The adiabatic theorem in \( \mathbb{C}P(1) \)

By using (1), we can write the Schrödinger equation explicitly in coordinates as

\[
\sum_{\nu} \hat{H}_\nu \psi^\nu = i \dot{\psi}^\mu.
\] (12)

For a two-level system (a single qubit), one can write

\[
\begin{bmatrix}
\hat{H}_0^0 \psi^0 + \hat{H}_0^1 \psi^1 = i \dot{\psi}^0 \\
\hat{H}_1^0 \psi^0 + \hat{H}_1^1 \psi^1 = i \dot{\psi}^1
\end{bmatrix}
\text{ with } \dot{H}(t) = \begin{pmatrix}
\hat{H}_0^0(t) & \hat{H}_0^1(t) \\
\hat{H}_1^0(t) & \hat{H}_1^1(t)
\end{pmatrix}.
\] (13)

We can project this equation of motion over the sphere \( \mathbb{C}P(1) \) (by using the complex projective coordinate \( \xi = \psi^1/\psi^0 \)) as a single complex-valued first-order differential equation over the space of rays:

\[
\dot{\xi} = i \left[ \xi (\hat{H}_0^0 - \hat{H}_1^1) + \xi^2 (\hat{H}_1^0 - \hat{H}_0^1) \right].
\] (14)

We find it useful to define the following real functions of time in terms of the matrix elements of the Hamiltonian:

\[
\Omega(t) = \hat{H}_0^0 - \hat{H}_1^1 \quad \text{and} \quad \hat{H}_0^0 = \overline{\hat{H}_0^0} = R(t) e^{-i \Omega(t)}.
\] (15)

Note that, though a generic \( 2 \times 2 \) Hermitian matrix has four independent matrix elements, the projection of the motion on \( \mathbb{C}P(1) \) needs at most three arbitrary time functions. This is because the projection of the motion on ray space is invariant under a translation-by-identity transformation as \( \dot{\hat{H}} \rightarrow \dot{\hat{H}} + f(t) \hat{I} \) since this would modify only the overall phase of the state vector but clearly leaves the value of \( \Omega \) unchanged. With these three defined parameters together with the stereographic coordinates provided by (2), we arrive at the following pair of coupled non-autonomous ODEs over the Bloch sphere:

\[
\begin{bmatrix}
\dot{\theta} = -2R(t) \sin(\varphi - \lambda(t)) \\
\dot{\varphi} = \Omega(t) - 2R(t) \cot \theta \cos(\varphi - \lambda(t))
\end{bmatrix}.
\] (16)

For an autonomous system, taking \( \{|u_0\}, \{|u_1\}\} \) as the energy eigenstates \( \hat{H} |u_0\rangle = E_0 |u_0\rangle \) of the Hamiltonian, the equation of motion simplifies drastically to

\[
\begin{bmatrix}
\dot{\theta} = 0 \\
\dot{\varphi} = \Omega = \text{constant}
\end{bmatrix}
\text{ with trivial solution } \begin{bmatrix}
\theta = \theta_0 \\
\varphi = \Omega t + \varphi_0
\end{bmatrix}.
\] (17)

where \( |u_0\rangle \) and \( |u_1\rangle \) represent, respectively, the north and south poles of the Bloch sphere (see figure 3). The possible time evolutions are the uniform circular motions around the pole-axis with constant angular velocity \( \Omega = E_0 - E_1 \). For an arbitrary non-autonomous system, we may diagonalize instantaneously the Hamiltonian \( \hat{H}(t) \), finding time-dependent energy eigenkets \( \{|E_0(t)\rangle, |E_1(t)\rangle\} \) represented by the antipodal points \( P_0(t) \) and \( P_1(t) \) on the Bloch sphere \( S^2 \). The motion is now given instantaneously by a circular motion around the instantaneous axis determined by \( P_0(t) \) and \( P_1(t) \) with instantaneous angular velocity \( \omega(t) = E_0(t) - E_1(t) \).

\[
\omega(t) = E_0(t) - E_1(t).
\] (18)
We may introduce coordinates $\theta'(t)$ and $\phi'(t)$ for the point $P_0(t)$ by the stereographic map
\[
\tan \left( \frac{\theta'}{2} \right) e^{i\phi'} = \left( \langle u^0 | E_0(t) \rangle \right)^{-1} \langle u^1 | E_0(t) \rangle.
\] (19)

The right side of the above equation can be computed by the eigenvalue equation:
\[
\sum_\sigma |E^\sigma|_{\mu_\sigma} (H_\sigma^0 - E_\sigma^0 \delta_\sigma^0) = 0,
\] (20)
giving us
\[
\begin{align*}
E^0 &= \frac{1}{2} \left[ H_0^0 + H_1^1 + (\Omega^2 + 4R^2)^{1/2} \right] \\
E^1 &= \frac{1}{2} \left[ H_0^0 + H_1^1 - (\Omega^2 + 4R^2)^{1/2} \right].
\end{align*}
\] (21)

We make now the following identifications:
\[
\begin{cases}
\lambda(t) = \phi'(t) \\
R(t) = \frac{1}{2} \omega(t) \sin \theta'(t) \\
\Omega(t) = \omega(t) \cos \theta'(t)
\end{cases}
\] (22)
and we arrive at a very convenient form for our coupled ODEs over the sphere $\mathbb{CP}(1)$:
\[
\begin{align*}
\dot{\theta} &= \omega(t) \sin \theta'(t) \sin (\phi - \phi'(t)) \\
\dot{\phi} &= \omega(t) \left[ -\cos \theta'(t) + \sin \theta'(t) \cot \theta \cos (\phi - \phi'(t)) \right].
\end{align*}
\] (23)

The three time-dependent parameters that describe our Hamiltonian can now be directly related to the geometry of the Riemann sphere: $\omega(t)$ gives us the instantaneous angular velocity of the motion in turn of the instantaneous axis of revolution whose direction is described by $(\theta'(t), \phi'(t))$. Surprisingly, this seemingly abstract formulation is actually implemented by nature as the motion of an electron spin immersed in a time-dependent classical magnetic field given by (see [21], for instance)
\[
\vec{B}(t) = \frac{m}{e} \omega(t) \left[ \sin \theta'(t) (\cos \phi(t) \hat{i} + \sin \phi(t) \hat{j}) + \cos \theta'(t) \hat{k} \right] 
\] (with $\hbar = c = 1$).

Let us consider now a very specific kind of motion where the Hamiltonian (the point $P_0(t)$) describes a uniform rotation on a great circle of the sphere which we choose to be the equator without any loss of generality:
\[
\theta'(t) = \pi/2 \quad \phi'(t) = \Omega t \quad \text{and} \quad \omega = \omega_0 = \text{constant}.
\]
Equation (23) then becomes
\[
\begin{align*}
\dot{\theta} &= \omega_0 \sin(\varphi - \Omega t) \\
\dot{\varphi} &= \omega_0 \cot \theta \cos(\varphi - \Omega t).
\end{align*}
\] (24)
We choose initial conditions
\[
\theta(0) = 0 \quad \text{and} \quad \varphi(0) = 0
\] (25)
to match the initial state vector as an energy eigenket. The above initial conditions together with equation (24) lead to
\[
\dot{\theta}(0) = 0 \quad \text{and} \quad \dot{\varphi}(0) = 0.
\] (26)
To prove the adiabatic theorem for this special case, all we need is to ensure that for a sufficiently small value of \(\omega_0/\Omega\), the conditions
\[
\theta(t) \approx \frac{\pi}{2} \quad \text{and} \quad \varphi(t) \approx \Omega t
\]
must hold in a self-consistent manner. Indeed, given that \(\varphi \approx \Omega t\), equation (24) implies that \(\theta \approx \arccot(\Omega/\omega_0) \approx \arccot(0) = \pi/2\). To prove the contrary implication, assume \(\theta \approx \pi/2\).
It is convenient then to introduce new coordinates
\[
\begin{align*}
\varepsilon(t) &= \theta(t) - \frac{\pi}{2} \\
\eta(t) &= 2(\varphi - \Omega t)
\end{align*}
\] (27)
which gives us
\[
\begin{align*}
\dot{\varepsilon} &= \omega_0 \sin(\eta/2) \\
\dot{\eta} &= -2[\Omega + \omega_0 \tan \varepsilon \cos(\eta/2)]
\end{align*}
\] with \(\varepsilon(0) = 0\) and \(\eta(0) = 0\), and \(\dot{\varepsilon}(0) = 0\) and \(\dot{\eta}(0) = -2\Omega\).
(28)
By taking the time derivative of the second equation (28) and using the first equation together with the condition \(\theta(t) \approx \pi/2 \implies \varepsilon \approx 0\), one obtains the approximate differential equation
\[
\ddot{\eta} \approx -\omega_0^2 \sin \eta
\] (29)
which we recognize immediately as the well-known simple pendulum equation (from undergraduate classical mechanics) with ‘conserved energy’ given by
\[
E = T + U = \frac{1}{2} \dot{\eta}^2 - \omega_0^2 \cos \eta.
\] (30)
It follows from the initial conditions in (28) that the ‘total mechanical energy’ of the pendulum is
\[
E = 2\Omega^2 - \omega_0^2
\] (31)
with minimum ‘potential energy’ at \(U(\eta_{\min}) = U(0) = -\omega_0^2\) (see figure 4). It is easy to see in this set up that one can confine the motion as close as one desires to \(\eta_{\min} = 0\) by making the ratio \(\omega_0/\Omega\) as large as necessary. So we find that the implication \(\theta \approx \pi/2 \implies \varphi \approx \Omega t\) must hold true. To extend the proof to an arbitrary path over the sphere, one must only need to remember that such path can be divided in \(N\) (let us say equal segments) sub-paths with \(N\) being sufficiently large so that each segment may be considered as approximately a geodesic path in such a way that the above argument can be applied piece-wise.
We have plotted (see figures 5 and 6) some numerical implementations of our simplified model (equation (24)) with different values of \(\omega_0/\Omega\).
This analysis shows convincingly that the projected state ‘whirls’ around the instantaneous north pole ever more tightly as the factor \(\omega_0/\Omega\) gets larger and larger.
3. Concluding remarks

We have briefly reviewed the geometry of quantum mechanics developed back in the 1980s by Aharonov and Anandan among others. We used these ideas to present an elementary geometric argument for the validity of the adiabatic theorem. We have used only the minimum geometric structure needed for the understanding of the adiabatic theorem, but there are many rich structures in the geometry of quantum mechanics. We refer the reader to [22–25] for more details on these matters.

Although we do not claim this as a complete proof, we are sure that this geometric reasoning can be easily implemented as a rigorous proof for those analytically inclined. We hope that this geometric picture is sufficiently straightforward to turn the concept of quantum adiabatic computation into a much more understandable tool for undergraduate students of physics and even for specialists.

The modern theories of quantum computation and information have become a central research field in physics and the quest for the best models of quantum computing (quantum gate model, adiabatic quantum computing, cluster or measurement-based quantum computation) and even to establish the equivalence (or not) between their true powers of computation (computability and complexity classes) are still open questions. A better intuitive
understanding and presentation of the key physical and mathematical structures behind their functioning seems to be a sensible enterprise. We believe we have presented the most direct and pedagogical framework so far of this important theorem, at least for two-level systems.

Acknowledgments

A C Lobo wishes to acknowledge financial support from NUPEC-Fundação Gorceix, C A Ribeiro, P R Dieguez and R A Ribeiro acknowledge financial support from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). The authors also wish to thank the anonymous referee for many valuable suggestions that helped to improve the quality of this manuscript. The authors thank Fernanda Lobo Bellehumeur for assistance with the English language and also thank D Brody for the references on the geometry of quantum mechanics.

References

[1] Born M and Fock V 1928 Z. Phys. A 51 165
[2] Kato T 1950 J. Phys. Soc. Japan 5 435
[3] Berry M 1984 Proc. R. Soc. Lond. A 392 45
[4] Aharonov Y and Anandan J 1987 Phys. Rev. Lett. 58 1593
[5] Farhi E, Goldstone J, Gutmann S and Sipser M 2000 arXiv:quant-ph/0001106
[6] Aharonov D, van Dam W, Kempe J, Landau Z, Lloyd S and Regev O 2004 SIAM Rev. 50 755
[7] Marzlin K and Sanders B 2004 Phys. Rev. Lett. 93 160408
[8] Wu Z and Yang H 2005 Phys. Rev. A 72 012114
[9] Tong D M, Singh K, Kwek L C and Oh C H 2005 Phys. Rev. Lett. 95 110407
[10] Ma J, Zhang Y, Wang E and Wu B 2006 Phys. Rev. Lett. 97 128902

Figure 6. Three plots for \( \omega_{0}/\Omega \) given by \( 10^1 \), \( 10^2 \) and \( 10^3 \).
[11] Wei Z and Ying M 2007 Phys. Rev. A 76 024304
[12] Amin M 2009 Phys. Rev. Lett. 102 220401
[13] Boixo S and Somma R 2010 Phys. Rev. A 81 032308
[14] Frasca M 2011 arXiv:1107.4971
[15] Rigolin G and Ortiz G 2011 arXiv:1111.5333
[16] Griffiths D 1995 Introduction to Quantum Mechanics (Englewood Cliffs, NJ: Prentice Hall)
[17] Messiah A 1999 Quantum Mechanics vol 1 and 2 (New York: Dover)
[18] Kasuga T 1961 Proc. Japan Acad. A 37 366
[19] Anandan J and Aharonov Y 1990 Phys. Rev. Lett. 65 1697
[20] Brody D 2003 J. Phys. A: Math. Gen. 36 5587
[21] Sakurai J J 1994 Modern Quantum Mechanics 2nd edn (Reading, MA: Addison-Wesley)
[22] Arnold V 1989 Mathematical Methods of Classical Mechanics vol 60 (Berlin: Springer)
[23] Kibble T 1979 Commun. Math. Phys. 65 189
[24] Page D 1987 Phys. Rev. A 36 3479
[25] Brody D and Hughston L 2001 J. Geom. Phys. 38 19