Maintaining a fixed observable after a control laser is turned off

Erik Anson

August 16, 2008

1 Introduction

The current methods of quantum control use shaped laser pulses as photonic reagents to alter the natural propagation of a system's wavefunction, driving the expectation value $\langle \mathbf{O} \rangle$ of an observable towards the desired value. This can be used to selectively ionize certain species within a plasma, align molecules, effect electronic transitions, or break bonds, among other uses.

However, all of this is in the presence of a strong electric field (i.e. the control pulse) which limits the ability to accurately probe the system or use it for further experiments. Once the control field is gone, the system's natural propagation quickly changes $\langle \mathbf{O} \rangle$, making it a difficult experimental problem to measure or otherwise utilize the system while it is still in its "controlled" state. The goal of this research was therefore to find target states that would not only satisfy the initial demands on $\langle \mathbf{O} \rangle$, but would also maintain that value (or guide it along a chosen path) for as long as possible after the control laser was turned off.

2 Mathematical Formulation of the problem

The expectation value of an observable is given by the equation

$$\langle \mathbf{O} \rangle = \langle \psi | \mathbf{O} | \psi \rangle \tag{1}$$

However, by Schrödinger's equation, the wavefunction $|\psi\rangle$ varies with time:

$$|\psi\rangle_t = e^{\frac{-i\mathbf{H}t}{\hbar}} |\psi\rangle_0 \tag{2}$$

where $\mathbf{H}$ is the Hamiltonian of the system once the control laser is off. If we expand this in the basis of the (normalized) energy eigenstates of the system, $\{\phi_n\}$, such that $\mathbf{H}|\phi_n\rangle = E_n|\phi_n\rangle$, then this becomes

$$|\psi\rangle_t = \sum_n a_n e^{\frac{-iE_n t}{\hbar}} |\phi_n\rangle \tag{3}$$

where the coefficients $a_n = \langle\psi|\phi_n\rangle$, and $\sum |a_n|^2 = 1$ because $|\psi\rangle$ must be normalized. Plugging this back into equation (1) and writing out each term of the product yields

$$\langle \mathbf{O}(t)\rangle = \sum_{nm} a_n a_m^* \langle\phi_m|\mathbf{O}|\phi_n\rangle e^{\frac{i(E_m - E_n)t}{\hbar}} \tag{4}$$

$$= \sum_{nm} a_n a_m^* \mathbf{O}_{mn} e^{\frac{i(E_m - E_n)t}{\hbar}} \tag{5}$$

where * denotes a complex conjugate. This equation describes how the expectation value of interest, $\langle\mathbf{O}\rangle$, varies with time. However, each of the matrix elements $\mathbf{O}_{mn}$ and energies E_n is a property of the system. The problem is therefore to optimally choose the a_n values to get as close as we can to the desired $\langle\mathbf{O}(t)\rangle$ values. Those a_n values define an optimal initial state of the system; the production of that state could then be posed as a more traditional control problem (allowing the use of control lasers). Although in reality $\mathbf{H}$ and $\mathbf{O}$ are infinite in size, for computational reasons I always approximate them with finite Hermitian matrices, which is equivalent to approximating $|\psi\rangle$ using a finite basis set of eigenstates.

3 Methods of Solution

The definition of "optimal" above has been left intentionally abstract, because different metrics of "quality" may be necessary for different applications, and may suggest different methods of solution. One such metric is to define a cost functional J, which measures how far $\langle\mathbf{O}(t)\rangle$ is from a chosen function $f(t)$ over some time interval of interest:

$$J \stackrel{\text{def}}{=} \int_{t=0}^{T} \left[\langle\mathbf{O}(t)\rangle - f(t)\right]^2 \, dt \tag{6}$$

This type of metric lends itself to gradient optimization, to find the a_n values that minimize J. Using a numerical integration with resolution $\Delta t = T \cdot 10^{-3}$, this gradient can be computed several times a second, leading to rapid convergence. If we choose a function $f(t)$ that is consistent with equation (5), then an exact solution exists where $J = 0$. In those cases, there do not appear to be any "traps," i.e. local minima that would prevent a gradient-based algorithm from finding the global minimum. This behavior did not change even when random "noise" was added to $f(t)$. However, if $f(t)$ is "un-physical," i.e. it is *not* consistent with equation (5), then traps often appear, making it difficult to get reliable results.

Alternatively, we can use as a metric the presence or absence of certain features of $\langle\mathbf{O}(t)\rangle$ that are important to a given experiment, e.g. certain chosen coefficients of its Fourier or Taylor series, or certain chosen function values at a series of time points. Each of these demands can be cast in the form of an observable, i.e. each demand is of the form $\langle\psi|\mathbf{M}_k|\psi\rangle = c_k$, for some

Hermitian matrix $\mathbf{M}_k$ and real scalar c_k. If these matrices are NxN, we can make a system of N such equations (one of which is $\langle\psi|\mathbf{I}|\psi\rangle = 1$, to ensure that $|psi\rangle$ is normalized). As long as the equations are consistent (i.e. as long as our demands are not impossible or mutually exclusive), we can then solve that system of equations iteratively, giving us an initial state which satisfies each of our demands.

If we pick any metric which can be calculated numerically, we can also choose to use a "genetic algorithm," which uses evolution as a model for optimization. In a genetic algorithm, there is a "population" of possible solutions (in this case, each member of the population is an initial state vector, $|\psi\rangle_0$), which evolves over time. For each "generation" of this population, the "fitness" (a metric of the degree to which it satisfies our demands, e.g. the J value above) of each member of the population is calculated. The least fit (e.g. the bottom 60%) are then removed, and those that remain "reproduce" in pairs, where the "offspring" are a random combination of the two "parents" (in this case, the offspring have some of the a_n values of one parent and some of the a_n values of the other). The population then undergoes random mutation, after which the process starts again. This approach is very useful, because its stochastic nature can avoid falling into the same traps as a gradient algorithm. However, it can have slow convergence, especially near the peak of the optimization landscape.

4 Findings

Above, the definition of "optimal" was more abstract, but let us consider the class of problems where we were trying to get $\langle\mathbf{O}(t)\rangle$ to approximate $f(t) = c$, where c is a constant of our choice. For this, we use the squared error metric J, as defined in equation (6).

Although the maximum and minimum values that $\langle\mathbf{O}\rangle$ can take on are defined by the extremal eigenvalues of $\mathbf{O}$ ($\lambda_{\max}$ and $\lambda_{\min}$), the values on the diagonal of $\mathbf{O}$ determine how well and for how long we can set $\langle\mathbf{O}(t)\rangle \approx c$ for a given c value. Consider the following cases:

- $c = \mathbf{O}_{kk}$ for some k

  We can let $|\psi\rangle_0 = |\phi_k\rangle$ (the k^{th} eigenstate), which yields

  $$\langle\mathbf{O}(t)\rangle = |a_k|^2\mathbf{O}_{kk}e^0 \tag{7}$$
  $$= \mathbf{O}_{kk} \tag{8}$$

  which is independent of time. In general, the k^{th} pure energy eigenstate yields an $\langle\mathbf{O}\rangle$ value that is constant at the k^{th} value on the main diagonal of $\mathbf{O}$.

- $\mathbf{O}_{mn} = 0$ for some n and m

  We can choose to limit our wavefunction to a superposition of only the n^{th} and m^{th} eigenstates, i.e. $|\psi\rangle_0 = a_n|\phi_n\rangle + a_m|\phi_m\rangle$, where $|a_n|^2 + |a_m|^2 = 1$

by normalization. We then have:

$$\langle \mathbf{O}(t) \rangle = |a_n|^2 \mathbf{O}_{nn} + |a_m|^2 \mathbf{O}_{mm} + 2 \cdot \mathrm{Re}\left[a_n a_m^* \mathbf{O}_{mn} e^{\frac{i(E_m - E_n)t}{\hbar}} \right] \quad (9)$$

$$= \mathbf{O}_{mm} + |a_n|^2 \left(\mathbf{O}_{nn} - \mathbf{O}_{mm} \right) \quad (10)$$

This expression is constant in time, and can be tuned to any value between $\mathbf{O}_{nn}$ and $\mathbf{O}_{mm}$ by choosing $|a_n|$.

- $E_n = E_m$ (i.e. $|\phi_n\rangle$ and $|\phi_m\rangle$ are degenerate eigenstates)
  Again, we let $|\psi\rangle = a_n|\phi_n\rangle + a_m|\phi_m\rangle$, where $|a_n|^2 + |a_m|^2 = 1$, yielding

$$\langle \mathbf{O}(t) \rangle = |a_n|^2 \mathbf{O}_{nn} + |a_m|^2 \mathbf{O}_{mm} + 2 \cdot \mathrm{Re}\left[a_n a_m^* \mathbf{O}_{mn} e^0 \right] \quad (11)$$

  Separation of the complex values a_n, a_m^*, and $\mathbf{O}_{mn}$ into their magnitude and phase components yields:

$$\langle \mathbf{O}(t) \rangle = \mathbf{O}_{mm} + |a_n|^2 \left(\mathbf{O}_{nn} - \mathbf{O}_{mm} \right) + 2 \cdot \mathrm{Re}\left[|a_n| e^{i\theta_n} |a_m| e^{-i\theta_m} |\mathbf{O}_{mn}| e^{i\theta_{mn}} \right] \quad (12)$$

$$= \mathbf{O}_{mm} + |a_n|^2 \left(\mathbf{O}_{nn} - \mathbf{O}_{mm} \right) + 2 \cdot |a_n||a_m||\mathbf{O}_{mn}| \cos(\theta_n - \theta_m + \theta_{mn}) \quad (13)$$

  This is again constant in time, and we can set θ_n and θ_m to any values we choose. If we choose to set $\theta_n - \theta_m + \theta_{mn} = \frac{\pi}{2}$, then we are back to the previous case, and we can set $\langle \mathbf{O}(t) \rangle$ to be any constant between $\mathbf{O}_{nn}$ and $\mathbf{O}_{mm}$. However, we can also reach values outside that range (to a degree that increases non-linearly with $|\mathbf{O}_{mn}|$) by manipulating the θ values.

Beyond these special cases, it is easier in general to set $\langle \mathbf{O}(t) \rangle \approx c$ if c is between the lowest and highest diagonal values of $\mathbf{O}$. If we separate the time-dependent and time-independent terms of equation (5), we get $\sum |a_n|^2 \mathbf{O}_{nn}$ (which is always between the highest and lowest diagonal values) plus sinusoidal oscillations in time. Therefore, it is impossible to set $\langle \mathbf{O}(t) \rangle \approx c$ for any c outside of that range without including non-trivial oscillatory terms (which must increase in magnitude as c gets further away from the nearest diagonal value). These large oscillations mean that $\langle \mathbf{O}(t) \rangle$ can't remain anywhere near a constant target value over long timescales.

As an additional result, I proved that if we want $\langle \mathbf{O} \rangle$ to reach its highest possible value (the highest eigenvalue of $\mathbf{O}$, $\lambda_{\max}$) at some time t_0, that constraint fully describes the wavefunction at all times. Moreover, $\langle \mathbf{O}(t) \rangle$ can not be held constant at $\lambda_{\max}$ for any non-zero length of time. Demanding that $\langle \mathbf{O}(t) \rangle$ come close to (but not reach) $\lambda_{\max}$ should also put strict limitations on its future behavior, but so far I have no quantitative results in that area.

Although I did most of my research on randomly-generated Hermitian matrices, there are some special cases that bear mentioning:

- $[\mathbf{H}_0, \mathbf{O}] = 0$

  If the Hamiltonian and observable matrices commute, then $\mathbf{O}$ is diagonal in the basis of the energy eigenstates. This means that equation (10) applies to any pair of diagonal values, since $\mathbf{O}_{mn} = 0$ for any choice of distinct m and n values. In addition, the diagonal values and the eigenvalues are the same. This means that (in the absence of a field) $\langle \mathbf{O}(t) \rangle = \langle \mathbf{O}(0) \rangle$, and so we can set c equal to any value between $\lambda_{\min}$ and $\lambda_{\max}$ and hold $\langle \mathbf{O} \rangle$ at that value indefinitely.

- Alignment

  One of the goals of laser-based control is to align linear systems along a chosen axis. For this, we consider the observable $\mathbf{O} = \cos^2(\theta)$, where θ is the angle between the symmetry axis of the system and the chosen alignment axis, and the system is approximated as a rigid rotor. For any rotational state $|J\rangle$, the only non-zero matrix elements off of the main diagonal are $\langle J|\mathbf{O}|J-2\rangle$ and $\langle J|\mathbf{O}|J+2\rangle$, meaning that equation (10) can be applied extensively. The lowest and highest diagonal values of $\mathbf{O}$ are $\mathbf{O}_{11} = 1/3$ and $\mathbf{O}_{22} = 0.6$, respectively, and $\mathbf{O}_{21} = 0$. This means that we can find an exact solution for $|\psi\rangle_0$ that satisfies $\langle \mathbf{O}(t) \rangle = c$ for any c between $1/3$ and 0.6. Interestingly, this solution requires only the two lowest energy states, making it an easy target if the system is sufficiently cold.

- Particle in a box

  Another special case of interest is a particle confined in a one-dimensional box of length L, with the position x as the observable. All of the diagonal values of $\mathbf{O}$ in this case equal $L/2$. $\langle \mathbf{O} \rangle$ is linearly dependent on any given coefficient a_n, but in most cases trying to adjust a second coefficient a_m at the same time introduces second-order terms. But in this case, $\mathbf{O}_{mn}$ is always zero if $m \neq n$ and $m - n$ is even. This means that if we hold the even-numbered a values constant, we can adjust any number of the odd a values without introducing second-order dependencies (i.e. there is no $|a_1||a_3|$ term, so changing the value of a_1 doesn't change the value of $\partial\langle \mathbf{O} \rangle / \partial a_3$). This makes the problem linear, except for the normalization requirement. However, the equations for the time derivatives of $\langle \mathbf{O} \rangle$ can be linearized in the same way, and so they can be set to zero by solving a system of linear equations (and values equal to zero are unaffected by normalization).

5 Conclusion

To some extent, an expectation value $\langle \mathbf{O} \rangle$ can be held near a chosen value c after the control laser has been turned off. However, the duration for which it can be held there decreases rapidly as c gets farther away from the nearest diagonal value (and closer to the extremal eigenvalues) of $\mathbf{O}$. Such demands are fulfilled

only by a very small fraction of all possible wavefunctions, which gives us less flexibility in attempting to fulfill additional demands. For example, if c is close to $\lambda_{\max}$, then $|\psi\rangle$ must be similar to the corresponding eigenstate of $\mathbf{O}$. This restriction makes it impossible to hold $\langle \mathbf{O} \rangle$ near c for any significant duration. However, for less demanding values of c, the methods listed here can find initial states which maintain that expectation value for much longer than a random wavefunction. By taking into account the propagation of the wavefunction after the control laser is turned off, instead of only considering the instantaneous value of $\langle \mathbf{O} \rangle$ right at the end of the laser pulse, it is therefore possible to extend the useful "lifetime" of that final state. Previously, the "controlled" state could only be measured or utilized nearly instantaneously after the control laser was turned off; however, this would relax that limitation, potentially allowing measurements and experiments that were previously impossible.

6 Further research

There are several avenues of further research which seem promising.

- Information Theory

  By generating random initial wavefunctions and propagating them, we can see that the states which reach extremely high and low $\langle \mathbf{O} \rangle$ values occur very infrequently. It may therefore be possible to cast the demand $\langle \mathbf{O}(t_0) \rangle = c$ in terms of its information content I, and relate our flexibility in choosing values for $\langle \mathbf{O}(t) \rangle$ for $t > t_0$ in terms of the remaining information entropy of the wavefunction. The key would be to find a way to calculate the "mutual information" of two or more demands of the form $\langle \mathbf{O}(t_0) \rangle = c_0$ and $\langle \mathbf{O}(t_1) \rangle = c_1$ in terms of the values c_0, c_1, and $t_1 - t_0$.

- Exploring further the relevance of $[\mathbf{H}, \mathbf{O}]$

  So far, the only result I have with respect to $[\mathbf{H}, \mathbf{O}]$ is the trivial case in which $[\mathbf{H}, \mathbf{O}] = 0$. However, it seems likely that either $[\mathbf{H}, \mathbf{O}]$, $[\mathbf{O}_0, \mathbf{O}_t]$ (where $\mathbf{O}_t$ is the observable at time t in the Heisenberg picture), or both are important to a full understanding of the underlying dynamics of this problem.

- Quantitative rules governing observed trends

  I have identified the trade-offs associated with optimal control in the field-free regime, and can produce rough estimates for how long a chosen c value can be held given the system parameters. However, a unifying mathematical expression that defines the limits of field-free control would increase the practical usefulness of this research, and may lead to new discoveries about the underlying principles.