
STATE-DEPENDENT FORCES IN COLD QUANTUM GASES

Christopher Billington

Submitted in total fulfilment of the requirements
of the degree of Doctor of Philosophy

Supervisory committee:

Prof Kristian Helmerston

Dr Lincoln Turner

Dr Russell Anderson



School of Physics and Astronomy
Monash University

April, 2018

rev: 122 (f0bd3bb3978d)
author: chrisjbillington
date: Tue Apr 10 17:50:58 2018 +1000
summary: Decoherence stuff

This page intentionally left blank

Contents

| | |
|--|-----------|
| Contents | i |
| 8 Hidden variables for semiclassical models with state-dependent forces | 1 |
| 8.0.1 Semiclassical models | 2 |
| 8.0.2 Stern–Gerlach separation and evaporative cooling | 3 |
| 8.1 Hidden variable theories | 6 |
| 8.2 Overview of method | 8 |
| 8.3 Implementation details | 8 |
| 8.3.1 Numerically evaluating Schrödinger theory | 8 |
| 8.3.2 Time-dependent formulation of Tully’s fewest-switches algorithm | 10 |
| 8.3.3 Distinguishing between spatial and temporal transitions in arbitrary hidden-variable theories | 15 |
| 8.3.4 Velocity correction and classically disallowed transitions . . . | 15 |
| 8.4 Decoherence | 16 |
| 8.4.1 Back action of position measurement on internal state | 17 |
| 8.4.2 Reframing as continuous weak measurement of internal state | 19 |
| 8.4.3 Form of the decoherence factor and the quantum Zeno effect | 20 |
| 8.4.4 Approximate Markovian decoherence | 22 |
| 8.4.5 Continuously spawned mean auxiliary trajectories | 23 |
| 8.5 Algorithm | 24 |
| 8.5.1 Further approximations | 24 |
| 8.6 results | 24 |
| 8.7 Discussion and conclusion | 24 |
| 8.7.1 Speculative aside: Monte-Carlo wavefunctions and Stochastic Schrödinger equations | 24 |
| 8.8 Approximate Markovian decoherence rate for separating wavepackets | 24 |
| 8.9 Choice of wavepacket size | 29 |
| References | 31 |

rev: 122 (f0bd3bb3978d)
author: chrisjbillington
date: Tue Apr 10 17:50:58 2018 +1000
summary: Decoherence stuff

This page intentionally left blank

rev: 122 (f0bd3bb3978d)
author: chrisjbillington
date: Tue Apr 10 17:50:58 2018 +1000
summary: Decoherence stuff

Hidden variables for semiclassical models with state-dependent forces

Hidden variable theories [CITE, CITE] are interpretations of quantum mechanics which posit that there are definite states underlying quantum wavefunctions, such that quantum indeterminacy is an illusion—an emergent phenomenon rather than a fundamental fact. The Bell inequality [CITE] proves that any such theory must be *nonlocal* in order to explain all the predictions of quantum mechanics, and perhaps in light of this, most physicists surveyed [CITE] do not believe that hidden variables underlie physical reality.

However, by framing quantum systems in classical terms, hidden variable theories can provide an excellent computational tool for *approximate* models of quantum systems, when it is reasonable to approximate some degrees of freedom as classical, yet other degrees of freedom need to be modelled quantum mechanically. Just as hidden variable theories have framed the quantum world in terms that are agreeable to the classical view of the world in the minds of some interpreters of quantum mechanics, so can they bridge the gap between a *simulated* quantum world and a *simulated* classical world coexisting in the same computer simulation.

In this chapter I describe what I call the ‘hidden variables semiclassical’ (HVSC) method: a method of combining quantum simulations with classical simulations, with hidden variables bridging the gap between the classical and quantum degrees of freedom. In this introduction I describe existing semiclassical methods, the manner in which their quantum and classical parts are typically coupled based on expectation values (which I’m calling the ‘Ehrenfest method’), and in which regimes this is inaccurate—namely the Stern–Gerlach experiment and similar situations in which considerable entanglement between motional and internal degrees of freedom develop.

In [SECREP THE NEXT SECTION] I give the technical definition of a hidden variable theory, and motivate the use of such a theory for coupling quantum and classical degrees of freedom in such a way that the a semiclassical model can be made to agree more closely than the Ehrenfest method with the underlying fully quantum model it is approximating, and ultimately, with experiment.

In section [SECREP THE SECTION AFTER THAT] I then go into the implications of welcoming a hidden variable into a semiclassical model, including additional required assumptions and approximations, and I derive the equations of motion for the model as a whole. Then in section [SECREP] I discuss the method further including its limitations.

Throughout this chapter, I make comparison between the HVSC and Ehrenfest methods. The two methods share most of the same flaws, save for HVSC’s ability to reproduce the Stern–Gerlach experiment. As such HVSC is an incremental improvement on the

rev: 122 (f0bd3bb3978d)
author: chrisjbillington
date: Tue Apr 10 17:50:58 2018 +1000
summary: Decoherence stuff

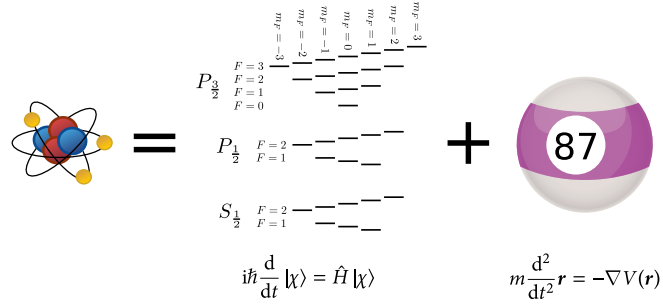


Figure 8.1: Artist's depiction of a semiclassical atom.

Ehrenfest method that solves this one specific problem. My method has a number of crude approximations—namely that wavepackets have a fixed size and that the decoherence of the atom's internal state is Markovian—and the wavepacket size and decoherence rate are estimated using approximate formulae. While crude, this is nonetheless an improvement over the Ehrenfest method's implicit assumption that decoherence does not occur at all (in which case the wavepacket size does not matter, as I will show), which is considerably less correct than even the crudest calculation of a Markovian decoherence rate. The HVSC method forces these assumptions to be made explicit, and serves as a minimal improvement to the Ehrenfest method, taking into account the simplest kind of decoherence only. Further improvements to the method are likely possible, but would be computationally more intensive to implement and unlikely to be worth it given the typical accuracy demanded of semiclassical simulations.

8.0.1 Semiclassical models

A semiclassical model is any model in which some degrees of freedom are treated quantum mechanically, and others classically. The most common combination is that of treating an atom's internal electronic state quantum mechanically and its motional degree of freedom classically. This is useful whenever the quantum effects of the atom's motion are not of interest, for example if temperatures are high and thus atomic wavelengths are short—such that quantum effects simply aren't visible in the motion of the particles and so they can accurately be modelled as classical billiard balls. The energy gaps between different electronic states are so large however that only at very high temperatures (at which atoms ionise anyway) do they start to appear as a continuum compared to thermal energy scales, and the interaction of different spin states of the atom with different optical and magnetic fields does not make them appear as classical continua either. Thus, quantum effects can be ignored for centre of mass motion of the atom, but not for the relative motion of its electrons with respect to the nucleus, or for the nuclear and electronic spin degrees of freedom.

In this regime, atoms are often modelled semiclassically, with these internal degrees of freedom modelled using a state vector $|\chi\rangle$ evolving according to a Hamiltonian \hat{H} via the Schrödinger equation, and the centre of mass motion modelled as a position \mathbf{r} and velocity \mathbf{v} evolving according to Newton's second law (Figure 8.1).

Once one has defined a potential function $V(\mathbf{r})$ and a Hamiltonian (also possibly varying with space) $\hat{H}(\mathbf{r})$, and other possible additions¹, ones job is done and the rest can be left to numerical differential equation solvers to evolve some concrete vector representation χ of $|\chi\rangle$ as well as the state variables \mathbf{r} and \mathbf{v} for motional degree of

¹Such using a Monte-Carlo wave-function method [CITE, SE-CREF?] to model the effect of spontaneous emission on $|\chi\rangle$, and modifying \mathbf{v} instantaneously by a random-direction recoil velocity upon each photon emission

freedom in time according to the coupled differential equations

$$\frac{d}{dt} |\chi\rangle = -\frac{i}{\hbar} \hat{H}(\mathbf{r}) |\chi\rangle, \quad (8.1)$$

$$\frac{d}{dt} \mathbf{v} = -\frac{1}{m} \nabla V(\mathbf{r}), \quad (8.2)$$

$$\frac{d}{dt} \mathbf{r} = \mathbf{v}. \quad (8.3)$$

In the next subsection I explain why it's not always that simple.

8.0.2 Stern–Gerlach separation and evaporative cooling

In the Stern–Gerlach experiment [CITE], particles with quantum spin and a magnetic moment—atoms for example—are fired as a beam through a region of space with a magnetic field gradient. The well known result is that two clusters of positions (if the atoms are spin $\frac{1}{2}$) are observed once the beam emerges, rather than a continuous smear of positions, indicating that angular momentum—like many quantities in quantum mechanics—is quantised.

This is the case even if one spin-polarises the particles before they are passed through the magnetic field gradient, say putting them in an eigenstate of the \hat{F}_x operator. Then, if the magnetic field is along the z direction, and the gradient is also in the z direction, two clusters of positions are also observed, even though all particles were in the same state when they entered the region in which there was a magnetic field gradient. This is a display of the indeterminacy of quantum mechanics: even though all particles had the same initial state, there were nonetheless different outcomes for each particle.

The Stern–Gerlach effect is a consequence of quantum mechanics, to be sure, but it has little to do with the wave nature of the atoms. If we introduced some double slits for the atoms to pass through in addition to the magnetic field gradient, then we would be seeing the wave nature of the atoms as interference patterns at the detection screen at the end of the experiment. But if we do not, and if the particles have short de-Broglie wavelengths, then quantum mechanics is not apparent in the motion of the particles through space—except via the influence of spin on seemingly choosing one trajectory or the other. The effect is well understood quantum mechanically, but is difficult to model semiclassically because even if we are happy to approximate wavepackets as small, the wavepackets do not take a single trajectory. Rather they split into two wavepackets, with the part of the superposition corresponding to one spin projection state (along the direction of the local magnetic field) moving one way, and the part of the superposition with the other spin projection going the other way. The trajectories can still be quite classical, it's just that there are two of them.

A similar situation exists in RF evaporative cooling (Section ??) of cold atoms en-route to BEC. Atoms are trapped in a magnetic trap, and are spin polarised so as to be fully spin down (for ^{87}Rb this is the trapped state) with respect to the local magnetic field at the position of each atom. The magnetic field's direction—not just its magnitude—varies in space, and so different atoms have different spins, but they are all spin down with respect to the quantisation axis of the local magnetic field.² As the atoms move through space, they move in orbits—punctuated by collisions—about the magnetic field zero at the centre of the trap, since they feel a force $F \propto -\nabla|B|$ due to the gradient of the Zeeman potential. Provided they are moving slowly (specifically, provided their Larmor precession period is short compared to the time the magnetic field as seen by the atom takes to change by a non-negligible fraction of its current value), the atoms' spins adiabatically follow the local field and remain spin-down, even as the field as seen by each atom fully reverses its direction every half orbital period.

²Out of habit we are already speaking semiclassically—to talk about the atoms' position is to assume they have one—and yet we're describing their spin quantum mechanically

Near the centre of the trap where the atoms are moving faster, the fields are small and therefore have large fractional derivatives and lead to large Larmor periods, adiabaticity no longer holds and the atoms may make spin transitions with respect to their local magnetic field. Once an atom passing close to the field zero has evolved into a superposition of spin projection states with respect to the local field, it is in a situation identical to the initial condition of the Stern–Gerlach experiment, causing the spin projection components to spatially separate in the magnetic field gradient. The spin up component is anti-trapped and repelled from the centre of the trap, and the zero spin projection component (since the groundstate of ^{87}Rb is spin-1) feels no force and moves in a straight line. The spin down component continues on an orbit about the field zero which is just as tight as before, unaffected by the close approach to the field zero other than being reduced in amplitude. Eventually a collision occurs, either with other atoms or with the walls of the vacuum system and the wavefunction collapses to choose one of these options, leading to atoms probabilistically leaving the trap (called Majorana losses [CITE]) or remaining trapped. Again, the trajectories can still be quite classical, it's just that there are three of them, and which is taken is probabilistic.

How can we model these effects semiclassically? Equations (8.1) to (8.3) are not sufficient, because there exists no single classical potential $V(\mathbf{r})$ that can describe the motion of the atoms. Rather, the atoms feel a different force depending on which spin state they are in. Just as the Hamiltonian can be a function of space, so can the potential be a function of the internal state of the atom: $V = V(\mathbf{r}, |\chi\rangle)$. Ehrenfest's theorem states that [CITE]

$$m \frac{d^2}{dt^2} \langle \hat{\mathbf{r}} \rangle = - \langle \nabla \hat{V} \rangle, \quad (8.4)$$

where the expectation values are over all degrees of freedom, not just motional. If we approximate a small wavepacket centred at the position \mathbf{r} in order to ignore the wave nature of the atoms, this becomes:

$$m \frac{d^2}{dt^2} \mathbf{r} = - \nabla \langle \chi | \hat{V}(\mathbf{r}) | \chi \rangle, \quad (8.5)$$

where the operator $\hat{V}(\mathbf{r})$ now only acts on the subspace of the internal state of the atom, since we have already taken an expectation value over (a small region of) space. Provided all potentials the atom is subjected to are included in the Hamiltonian for its internal state (including any energy offsets that do not depend explicitly on the internal state), this is nothing but

$$m \frac{d^2}{dt^2} \mathbf{r} = - \nabla \langle \chi | \hat{H}(\mathbf{r}) | \chi \rangle, \quad (8.6)$$

where \hat{H} is the Hamiltonian describing the evolution of the atom's internal state. We now can construct the *Ehrenfest semiclassical method* describing how the *expectation value* of a well localised atom's position evolves with time:

$$\frac{d}{dt} |\chi\rangle = -\frac{i}{\hbar} \hat{H}(\mathbf{r}) |\chi\rangle, \quad (8.7)$$

$$\frac{d}{dt} \mathbf{v} = -\frac{1}{m} \langle \chi | \hat{H}(\mathbf{r}) | \chi \rangle, \quad (8.8)$$

$$\frac{d}{dt} \mathbf{r} = \mathbf{v}. \quad (8.9)$$

The Ehrenfest semiclassical method is the same as the simple semiclassical method (8.1) to (8.3), except that it has an answer to the question “What should we use for $V(\mathbf{r})$ when

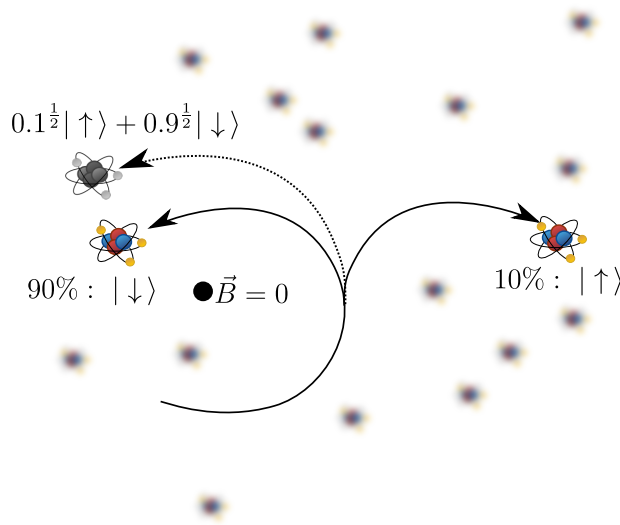


Figure 8.2: todo

the atom is in a superposition of states that feel different potentials?”, which is “use the expectation value”.

This is all well and good if the expectation value of position is a good approximation to the situation being modelled. But in the Stern–Gerlach experiment and Majorana losses in magnetic traps, the expectation value of position is a poor match to reality. In the Stern–Gerlach experiment beginning with spin-polarised atoms, this would result in a single blob in the middle of the screen, rather than two. And in an atom approaching the field zero in a magnetic trap, it would result in the atom broadening its orbit somewhat, rather than splitting into multiple possible trajectories (Figure 8.2). Semiclassical simulations of evaporative cooling performed by Christopher Watkins [CITE] displayed an unphysical heating of the atom cloud that I believe is due to the Ehrenfest method’s inability to model Stern–Gerlach separation. In a real magnetic trap during evaporative cooling to Bose–Einstein condensation, the mean free path is large enough that the part of the wavepackets that are no longer trapped will usually leave the trap without colliding with any other atoms. This means that the energy the untrapped and anti-trapped components have gained relative to the trapped component moving away from the field zero is not usually shared with other atoms upon collision—the extra energy leaves with the atoms. However, if a close approach to the magnetic field zero merely means a broadening of the atoms orbit, then the extra energy does not leave as fast, if at all, and can be shared with other atoms via collisions, turning what would have been an atom loss effect into an overall gradual heating of the cloud.

So how can we modify a semiclassical method to choose only one trajectory? Firstly, since each trajectory corresponds to one of the internal states (though some might be degenerate), our model must choose an internal state and use the classical trajectory corresponding to that internal state only. Secondly, since all atoms begin in identical states and yet some take one trajectory and some another, this choice must be probabilistic. Finally, the probabilities must be consistent with those from quantum mechanics, i.e. the Born rule: the probability of an atom taking each trajectory must be proportional to the squared amplitude of the internal state of the atom, projected onto the eigenstate corresponding to that trajectory.

rev: 122 (f0bd3bb3978d)
author: chrisjbillington
date: Tue Apr 10 17:50:58 2018 +1000
summary: Decoherence stuff

There exists a category of theories dealing with precisely this question of how to choose a specific state of a quantum system in a stochastic way, such that the probability of having chosen a state is equal to that given by quantum mechanics. Such theories are called “hidden variable” theories, and the number specifying which state has been chosen is the “hidden variable”.

8.1 Hidden variable theories

In his paper *Quantum computing and dynamical quantum models*, Aaronson defines a “dynamical quantum model”[CITE]:

A *dynamical quantum model* assigns an eigenstate to a specified observable even when no measurement is made, and gives a stochastic evolution rule for that eigenstate. Such a model yields a distribution of classical histories of a quantum state

In a later paper, *Quantum computing and hidden variables*, superseding the first, Aaronson renames dynamical quantum models *hidden-variable theories* and instead defines:

For us, a hidden-variable theory is simply a way to convert a unitary matrix that maps one quantum state to another, into a stochastic matrix that maps the initial probability distribution to the final one in some fixed basis

This is what a hidden-variable theory is for the purposes of this chapter as well. Though the second paper supersedes the first, the definition it gives is more tangible for us—we wish to *assign an eigenstate* to our atoms (choose one of the internal states in order to decide which trajectory to follow), and have a *stochastic evolution rule* for which eigenstate is chosen at any one time (allow the atom to begin taking a different trajectory if it makes transitions between states).

But the second definition is more specific. The stochastic evolution rule is in the form of a stochastic matrix, with elements equal to transition probabilities for some time interval. And the rule should be in some way based on the unitary evolution that the quantum system evolves according to in the same interval of time, such that the initial and final probabilities of the stochastic matrix and unitary evolution agree. That is, if a quantum state $|\chi\rangle$ evolves in a certain basis $\{|n\rangle\}$ according to a unitary $\hat{U}(t', t)$:

$$\begin{bmatrix} \chi_1(t') \\ \chi_2(t') \\ \chi_3(t') \\ \vdots \end{bmatrix} = \begin{bmatrix} U_{11}(t', t) & U_{12}(t', t) & U_{13}(t', t) & \dots \\ U_{21}(t', t) & U_{22}(t', t) & U_{23}(t', t) & \dots \\ U_{31}(t', t) & U_{32}(t', t) & U_{33}(t', t) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \chi_1(t) \\ \chi_2(t) \\ \chi_3(t) \\ \vdots \end{bmatrix}, \quad (8.10)$$

where $\chi_n = \langle n|\chi\rangle$ and $U_{nm}(t', t) = \langle n|\hat{U}(t', t)|m\rangle$, then a hidden variable theory is a matrix-valued function $S(U(t', t), \chi(t))$, where $\chi(t)$ and $U(t', t)$ are the vector and matrix representations of the state vector $|\chi\rangle(t)$ and unitary $\hat{U}(t', t)$ in the $\{|n\rangle\}$ basis, that satisfies:

$$\begin{bmatrix} |\chi_1(t')|^2 \\ |\chi_2(t')|^2 \\ |\chi_3(t')|^2 \\ \vdots \end{bmatrix} = \begin{bmatrix} S_{11}(t', t) & S_{12}(t', t) & S_{13}(t', t) & \dots \\ S_{21}(t', t) & S_{22}(t', t) & S_{23}(t', t) & \dots \\ S_{31}(t', t) & S_{32}(t', t) & S_{33}(t', t) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} |\chi_1(t)|^2 \\ |\chi_2(t)|^2 \\ |\chi_3(t)|^2 \\ \vdots \end{bmatrix}, \quad (8.11)$$

In essence, if quantum unitary evolution takes amplitudes to amplitudes, a hidden variable theory takes probabilities to probabilities. Thus the elements of S are conditional probabilities—or transition probabilities—for a hidden variable $\eta(t)$, giving the chance

that the state $|\eta\rangle \in \{|n\rangle\}$ assigned by the hidden variable will change from one to another in the given time interval: [double check indices]

$$\Pr(\eta(t')=i|\eta(t)=j) = S(U(t', t), \chi(t))_{ij}. \quad (8.12)$$

There are many ways to define functions S that satisfy this condition. The simplest is to ignore the unitary completely and set $S_{ij} = |\chi_i(t')|^2$ for all n , which yields:

$$\Pr(\eta(t')=i|\eta(t)=j) = |\chi_i(t')|^2, \quad (8.13)$$

that is that the hidden variable η is equally likely to transition to a given value regardless of its previous value, and regardless of the unitary, and that that the conditional transition probability from all input states is just the squared amplitude of the final state. This theory represents a hidden variable that will just jump between states randomly based on their amplitude, with no regard for whether there were actually amplitude flows between the states between which it is transitioning. Nonetheless, it matches the definition of a hidden variable theory— S is a stochastic matrix, and the hidden variable on average will spend an amount of time in each state consistent with the Born rule.

Aaronson outlines [CITE] some additional axioms that hidden variables ought to satisfy, if we are to take them seriously, including reasonable statements about symmetries, and insensitivity to small perturbations. He goes on to prove that the properties cannot all be satisfied simultaneously, but some theories satisfy more of them than others.

It is convenient to introduce the matrix P of absolute transition probabilities. Summing (8.12) over all possible initial values of the hidden variable yields the absolute probability of it having a particular final value after the given time interval:

$$\Pr(\eta(t')=i) = \sum_j \Pr(\eta(t')=i|\eta(t)=j) \Pr(\eta(t)=j), \quad (8.14)$$

which, recognising that the final and initial probabilities must be the final and initial squared amplitudes of the state vector, can be rewritten:

$$|\chi_i(t')|^2 = \sum_j S(U(t', t), \chi(t))_{ij} |\chi_j(t)|^2. \quad (8.15)$$

We now define the matrix of absolute transition probabilities P :

$$P_{ij} = S_{ij}(U(t', t), \chi(t)) |\chi_j(t)|^2, \quad (8.16)$$

such that:

$$|\chi_i(t')|^2 = \sum_j P_{ij}. \quad (8.17)$$

So we have that P has row sums equal to the final squared amplitudes. And, because $P_{ij} = S_{ij} |\chi_j(t)|^2$ and S is a stochastic matrix with column sums equal to one,³ we have that P must have row sums equal to the initial squared amplitudes.

Now that we have introduced P and shown what its row and column sums must be, we come to a particularly simply defined hidden variable theory called the Schrödinger theory, discussed in Aaronson's hidden-variables paper [CITE]. The idea is to form P by staring with the matrix of absolute values of U , and simply scaling its rows and columns to have the correct values:⁴

$$P = \begin{bmatrix} a_1 b_1 |U_{11}(t', t)| & a_1 b_2 |U_{12}(t', t)| & a_1 b_3 |U_{13}(t', t)| & \dots \\ a_2 b_1 |U_{21}(t', t)| & a_2 b_2 |U_{22}(t', t)| & a_2 b_3 |U_{23}(t', t)| & \dots \\ a_3 b_1 |U_{31}(t', t)| & a_3 b_2 |U_{32}(t', t)| & a_3 b_3 |U_{33}(t', t)| & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad (8.18)$$

rev: 122 (f0bd3bb3978d)
author: chrisjbillington
date: Tue Apr 10 17:50:58 2018 +1000
summary: Decoherence stuff

³In order to more clearly compare to quantum evolution, we are using *left stochastic* matrices, which multiply column vectors of probabilities from the left, contrary to the most common convention for stochastic matrices which is to multiply row matrices from the right. Thus S has unit column sums, whereas the corresponding right stochastic matrix (its transpose) would have unit row sums.

⁴One might expect that the absolute value squared might be a better choice, since this matches Fermi's golden rule. Aaronson argues in his paper, when discussing his own hidden-variable theory *Flow theory* why absolute values of elements of the unitary have nice properties, and we show later that the row and column scalings reproduce Fermi's golden rule in any case, as they must for the hidden variable

that is,

$$P_{ij} = a_i b_j |U_{ij}(t', t)|, \quad (8.19)$$

where the row scalings $\{a_i\}$ and column scalings $\{b_j\}$ satisfy:

$$\sum_j a_i b_j |U_{ij}(t', t)|^2 = |\chi_i(t')|^2, \quad (8.20)$$

$$\sum_i a_i b_j |U_{ij}(t', t)|^2 = |\chi_j(t)|^2, \quad (8.21)$$

which can be solved numerically, and then the Schrodinger theory stochastic matrix then able to be extracted by inverting (8.16):

$$S_{ij}(U(t', t), \chi(t)) = a_i b_j \frac{|U_{ij}(t', t)|}{|\chi_j(t)|^2}. \quad (8.22)$$

8.2 Overview of method

Summarise the method in high level terms, flag that details are to come, with a specific algorithm presented after the details.

8.3 Implementation details

8.3.1 Numerically evaluating Schrödinger theory

There are many ways to solve numerically for the row and column scalings, but there is one unique solution for the resulting scaled matrix.⁵ The simplest method is to simply alternate between scaling the rows to get the right row sums, then scaling the columns to get the right column sums, and repeating, that is, alternating between solving equation (8.20) for all a_i , and solving (8.21) for all b_j , until the result converges. This is called the Sinkhorn-Knopp method of r-c (row-column) scaling [CITE], but is computationally intensive, with slow convergence [CITE LINIAL]. An alternative is the method by Linial et al [CITE], which converges much faster. Both are iterative methods, and so in practice one can save the resulting row and column scalings at each integration step of a simulation and use them as the initial guesses for the same computation at the next integration step,⁶ providing considerable speedup.

For the case of a 2×2 system, the row and column scalings can be found analytically, with the result

$$a_1 = 1 \quad (8.23)$$

$$a_2 : a_2^2 + \left(\frac{1}{AB} - \frac{A}{B} |\chi_1(t)|^2 - \frac{B}{A} |\chi_2(t)|^2 \right) a_2 + \frac{|\chi_2(t')|^2}{|\chi_1(t')|^2} = 0 \quad (8.24)$$

$$b_1 = \frac{|\chi_1(t')|^2}{A + Ba_2} \quad (8.25)$$

$$b_2 = \frac{|\chi_2(t')|^2}{B + Aa_2}, \quad (8.26)$$

where a_2 is the positive solution to the given quadratic, $A = |U_{11}| = |U_{22}|$ and $B = |U_{12}| = |U_{21}|$. This result only holds in the case of Schrödinger theory for a state vector with two components subject to unitary evolution—not for row-column scaling in general—since it makes use of symmetries of unitary matrices and the fact that the initial

⁵The values of $\{a_i\}$ and $\{b_j\}$ are only determined up to an overall multiplication of each $\{a_i\}$ by a constant and division of each $\{b_j\}$ by the same constant, since only products $a_i b_j$ appear in the resulting scaled matrix

⁶With the caveat that since (as mentioned in the previous footnote) the row and column scalings are only determined up to an overall multiplication/division, occasional multiplication of all $\{a_i\}$ and division of all $\{b_j\}$ by a constant may be necessary to prevent the values numerically overflowing or underflowing in the middle of a simulation.

and final probabilities given by the squared amplitudes of the state vector components must sum to unity.

Some further notes on numerics: the above expressions for Schrödinger theory and its analytic expression for a 2×2 system involve dividing by elements of the unitary, and state populations, both of which may be zero or very close to zero. Whilst the relevant limits may exist, we cannot easily compute them numerically, and so I have taken to simply replacing small values of $|U_{ij}|$, $|\chi_i(t')|^2$ and $|\chi_j(t)|^2$ with a small constant $\varepsilon = 10^{-10}$, ensuring that the convergence criterion (which represents a tolerance for the sum squared error in the column sums) I pass to Linial's method is larger than the square of this by some margin, so as to allow convergence even though modifying the matrix elements may make the matrix no longer row-column scalable to higher precisions. A convergence criterion of 10^{-16} for Linial's algorithm allows a sufficient margin and implies the root sum squared error in column sums will be at most 10^{-8} , which is small compared to one—which is the sum of all column sums for a perfectly scaled matrix given that the column sums are probabilities that must add to unity.

Potentially faster algorithms exist for row-column scaling of matrices,⁷ for example, ones that treat the problem as a root finding problem or optimisation problem aimed at solving the simultaneous equations (8.20) and (8.21) or minimising the residuals [A FAST ALGORITHM FOR MATRIX BALANCING PHILIP A. KNIGHT AND DANIEL RUIZ]. I had initially great success with Newton's method for finding a root to this set of equations (after fixing $a_1 = 1$ to make them fully determined), which required considerably fewer iterations than Linial's method for small (3×3) random unitary matrices and random state vectors. However, the unitaries and state vectors in quantum mechanics are not random, and the fact that most elements of the unitary and state vector are zero when there is no evolution and the atom is in an eigenstate resulted in numerical difficulties with Newton's method that Linial's method does not seem to encounter. Similar to many of the methods in [A FAST ALGORITHM FOR MATRIX BALANCING], one could construct a hybrid method that takes a Newton step, and then checks the row sum and column sum residuals, and if they got larger compared to the previous step, ignores that step and takes a step of Linial's method instead. I have not attempted this, and for the moment use Linial's method, which I provide Cython code for in [APPENDIX].

A final note is that my hidden variables semiclassical method is not just agnostic to which matrix scaling algorithm is used, but that it is also not married to any particular hidden variable theory. An early version of the method [CITE THE ARXIV PAPER] was limited to two-component systems, and the probability of transition was computed as:

$$\Pr(\eta(t')=2|\eta(t)=1) = \max\left(0, |\chi_2(t')|^2 - |\chi_2(t)|^2\right), \quad (8.27)$$

$$\Pr(\eta(t')=1|\eta(t)=2) = \max\left(0, |\chi_1(t')|^2 - |\chi_1(t)|^2\right), \quad (8.28)$$

that is, I simply inspected the populations each step and declared any positive change in population of a state as a probability of transition from the other state. Since there were only two states, the originating state of the transition was unambiguous, but the method did not generalise to systems with three or more states.⁸ However, it resulted in final populations in simulations on average in agreement with the underlying Schrödinger wave equation, leading me to suspect that the exact hidden variable theory used is not crucial, so long as it satisfies a few of Aaronson's axioms so as not to behave like the product theory. I use Schrödinger theory fairly arbitrarily, it being the one that seemed most easily computable out of the two presented in Aaronson's paper [CITE], but one might try using Aaronson's flow theory, or inventing another altogether. However, the main conclusion of Aaronson's paper is that if we could know the entire history of a hidden variable, we could use it to make a computer more powerful than a quantum

⁷In my simulations for realistic 3×3 Unitaries corresponding to evolution over small time intervals, Linial's method converges to the aforementioned tolerance in about 100 – 200 iterations, taking about 20 – 40 μ s per matrix on my computer. So it could be improved upon, but it is not prohibitive

⁸This was before I coincidentally came across the definition of a stochastic hidden-variable theory in Aaronson's book *Quantum computing since Democritus* [CITE] and realised that what I had made was a hidden variable theory, allowing me to choose a more general one from his paper.

computer. It is therefore perhaps not surprising that hidden variable theories ought to be computationally expensive to simulate. One cause for optimism however, if one wants to speed up my hidden variables semiclassical method, is the possibility that Aaronson's conclusion relies on entanglement between systems. If it does, such that hidden variables of unentangled systems are not equivalent to super-quantum computers, then there is hope, because my method does not require the hidden variable to model entanglement (or even mixed states for that matter), and therefore it might be possible that a simpler hidden variable theory could be constructed that is sufficient for the purposes of my model, but is still fast to compute.

8.3.2 Time-dependent formulation of Tully's fewest-switches algorithm

[TODO flag this as speculative and untested unless I test it]

In this subsection I derive Tully's fewest-switches algorithm [CITE], with the distinction that the Hamiltonian may have arbitrary time-dependence. As such, the non-adiabatic transitions that can occur can be due to either the spatial variation of the Hamiltonian, or its time variation. It is important to be able to distinguish which of the two non-adiabatic effects is responsible for a given transition of the hidden variable, as energy conservation only applies in the case of a nonadiabatic transition due to spatial variation of the Hamiltonian, in which case the particle is paying/receiving the energy cost of a transition using its kinetic energy. However for a transition due to temporal variation of the Hamiltonian, energy can be added and removed from the particle by the driving field without conserving its total energy. Thus, when a time dependent potential is used in a hidden-variable semiclassical simulation, velocity corrections ought to only be performed following a hidden-variable transition if that transition was due to spatial variation in the Hamiltonian, and not temporal variation.

We begin with the time dependent Schrödinger equation for the internal state $|\chi(t)\rangle$ of an atom at position \mathbf{r} :

$$i\hbar \frac{d}{dt} |\chi(t)\rangle = \hat{H}(\mathbf{r}, t) |\chi(t)\rangle \quad (8.29)$$

Now we take the unitary $\hat{U}_H(\mathbf{r}, t)$ that transforms state vectors into the eigenbasis of \hat{H} such that

$$\hat{H}(\mathbf{r}, t) = \hat{U}_H^\dagger(\mathbf{r}, t) \hat{V}(\mathbf{r}, t) \hat{U}_H(\mathbf{r}, t), \quad (8.30)$$

where $\hat{V}(\mathbf{r}, t)$ is a diagonal operator with diagonals equal to the adiabatic potentials that each eigenstate of \hat{H} is subject to in the adiabatic approximation. We can therefore define the state vector in the eigenbasis of \hat{H} as⁹

$$|\chi_H(t)\rangle = \hat{U}_H(\mathbf{r}, t) |\chi(t)\rangle \quad (8.31)$$

$$\Rightarrow |\chi(t)\rangle = \hat{U}_H^\dagger(\mathbf{r}, t) |\chi_H(t)\rangle. \quad (8.32)$$

Substituting (8.31) into (8.29) and premultiplying by $\hat{U}(\mathbf{r}, t)$ yields

$$i\hbar \hat{U}(\mathbf{r}, t) \frac{d}{dt} \hat{U}_H^\dagger(\mathbf{r}, t) |\chi_H(t)\rangle = \hat{U}(\mathbf{r}, t) \hat{H}(\mathbf{r}, t) \hat{U}_H^\dagger(\mathbf{r}, t) |\chi_H(t)\rangle \quad (8.33)$$

which via the product rule and our definition of $\hat{V}(\mathbf{r}, t)$ simplifies to the differential equation obeyed by the transformed state vector $|\chi_H(t)\rangle$:

$$i\hbar \frac{d}{dt} |\chi_H(t)\rangle = \left[\hat{V}(\mathbf{r}, t) - i\hbar \hat{U}(\mathbf{r}, t) \frac{d}{dt} \hat{U}_H^\dagger(\mathbf{r}, t) \right] |\chi_H(t)\rangle \quad (8.34)$$

$$= \hat{H}_{\text{eff}} |\chi_H(t)\rangle. \quad (8.35)$$

⁹This is very similar to an interaction picture state vector [SEE SECREP], but as I have previously used the definition of an interaction picture as the transformation that diagonalises a *time-independent* Hamiltonian, this potentially time-dependent Hamiltonian does not satisfy the definition.

This equation has the same form as the Schrödinger equation, with the contents of the brackets comprising an effective Hamiltonian dictating the dynamics of the state vector in our chosen basis. Like a non-inertial reference frame in classical mechanics, use of this transformed basis has resulted in the appearance of an extra term in the Hamiltonian, the non-adiabatic coupling term depending on the time derivative of the transformation \hat{U}^\dagger .

Here we differ from Tully by proceeding without assuming that \hat{U} has no explicit time dependence. The total time derivative of \hat{U}^\dagger includes both its direct time dependence and the effect of motion through space; the latter obtainable via the chain rule:

$$\frac{d}{dt} \hat{U}_H^\dagger(\mathbf{r}, t) = \frac{\partial}{\partial t} \hat{U}_H^\dagger(\mathbf{r}, t) + \mathbf{v} \cdot \nabla \hat{U}_H^\dagger(\mathbf{r}, t), \quad (8.36)$$

where $\mathbf{v} = \frac{d\mathbf{r}}{dt}$. Thus (8.34) becomes:

$$i\hbar \frac{d}{dt} |\chi_H(t)\rangle = \left[\hat{V}(\mathbf{r}, t) - i\hbar \hat{U}_H(\mathbf{r}, t) \frac{\partial}{\partial t} \hat{U}_H^\dagger(\mathbf{r}, t) - i\hbar \mathbf{v} \cdot \hat{U}_H(\mathbf{r}, t) \nabla \hat{U}_H^\dagger(\mathbf{r}, t) \right] |\chi_H(t)\rangle. \quad (8.37)$$

The final term in brackets is identical to the non-adiabatic coupling term in the equation of motion as usually written in the surface-hopping literature [CITE SOME THINGS] being a matrix with elements (in the eigenbasis):

$$\left(-i\hbar \mathbf{v} \cdot U_H(\mathbf{r}, t) \nabla U_H^\dagger(\mathbf{r}, t) \right)_{ij} = -i\hbar \mathbf{v} \cdot \langle \chi_i(\mathbf{r}, t) | \nabla | \chi_j(\mathbf{r}, t) \rangle, \quad (8.38)$$

where $U_H(\mathbf{r}, t)$ is the matrix representation of U in any fixed basis (i.e, not the eigenbasis), $|\chi_i(\mathbf{r}, t)\rangle$ is the i^{th} eigenvector of $\hat{H}(\mathbf{r}, t)$, and $\langle \chi_i(\mathbf{r}, t) | \nabla | \chi_j(\mathbf{r}, t) \rangle$ is the *non-adiabatic coupling vector* between the i^{th} and j^{th} states referred to in the literature. The second to last term in brackets in (8.37) is the additional contribution due to the time-dependence of the Hamiltonian (more specifically: the time-dependence of its eigenbasis).

We now proceed identically to Tully, computing the rate of change of a an eigenstates population $|c_i(t)|^2$ as:

$$\frac{d}{dt} |c_i(t)|^2 = c_i(t) \frac{d}{dt} c_i^*(t) + c_i^*(t) \frac{d}{dt} c_i(t), \quad (8.39)$$

where via (8.37) we have:

$$\frac{d}{dt} c_i(t) = -\frac{i}{\hbar} \sum_j (H_{\text{eff}}(\mathbf{r}, t))_{ij} c_j(t) \quad (8.40)$$

$$\Rightarrow \frac{d}{dt} c_i(t) = -\frac{i}{\hbar} \sum_j \left[V_{ij}(\mathbf{r}, t) - i\hbar \langle \chi_i(\mathbf{r}, t) | (\partial_t + \mathbf{v} \cdot \nabla) | \chi_j(\mathbf{r}, t) \rangle \right] c_j(t), \quad (8.41)$$

This gives for the time rate of change of the population $|c_i(t)|^2$:

$$\frac{d}{dt} |c_i(t)|^2 = \left[-\frac{i}{\hbar} \sum_j c_i^*(t) (H_{\text{eff}}(\mathbf{r}, t))_{ij} c_j(t) \right] + \text{c.c.} \quad (8.42)$$

$$= -\frac{2}{\hbar} \sum_j \text{Im} \left(c_i^*(t) (H_{\text{eff}}(\mathbf{r}, t))_{ij} c_j(t) \right) \quad (8.43)$$

$$= -\frac{2}{\hbar} \sum_j \text{Im} \left(c_i^*(t) \left[V_{ij}(\mathbf{r}, t) - i\hbar \langle \chi_i(\mathbf{r}, t) | (\partial_t + \mathbf{v} \cdot \nabla) | \chi_j(\mathbf{r}, t) \rangle \right] c_j(t) \right). \quad (8.44)$$

¹⁰This is not always assumed in the surface hopping literature, since additional couplings are sometimes included in V which have not been removed by diagonalisation.

Since $V(\mathbf{r}, t)$ is diagonal¹⁰ and real, $c_i^*(t)V_{ij}(\mathbf{r}, t)c_j(t)$ is zero when $i \neq j$, and has no imaginary part when $i = j$, leaving us with

$$\frac{d}{dt}|c_i(t)|^2 = 2 \sum_j \text{Re} \left(c_i^*(t) \langle \chi_i(\mathbf{r}, t) | (\partial_t + \mathbf{v} \cdot \nabla) | \chi_j(\mathbf{r}, t) \rangle c_j(t) \right). \quad (8.45)$$

The change in $|c_i(t)|^2$ in a small interval dt is then:

$$|c_i(t+dt)|^2 - |c_i(t)|^2 = 2dt \sum_j \text{Re} \left(c_i^*(t) \langle \chi_i(\mathbf{r}, t) | (\partial_t + \mathbf{v} \cdot \nabla) | \chi_j(\mathbf{r}, t) \rangle c_j(t) \right). \quad (8.46)$$

This is the change in the probability of the particle being in the i^{th} state during that time interval. Tully identifies each term in the sum as a probability flow between a pair of states, and if non-negative, equates each term with the (unconditional) probability of a transition from the j^{th} state to the i^{th} state. We do the same, except that we identify two transition probabilities for each originating state, one due to the spatial variation in the eigenbasis, and one due to the temporal variation. To ensure we don't violate the criterion that on a two-state basis only the minimum number of hops consistent with the total probability flow occur, we clip each probability from above to the probability of any transition occurring at all. This gives us transition probability matrix elements

$$P_{ij}^{\text{space}} = \min \{ q_{ij}^{\text{total}}, q_{ij}^{\text{space}} \}, \quad (8.47)$$

$$P_{ij}^{\text{time}} = \min \{ q_{ij}^{\text{total}}, q_{ij}^{\text{time}} \}, \quad (8.48)$$

where

$$q_{ij}^{\text{space}} = 2dt \text{Re} \left(c_i^*(t) \langle \chi_i(\mathbf{r}, t) | \mathbf{v} \cdot \nabla | \chi_j(\mathbf{r}, t) \rangle c_j(t) \right), \quad (8.49)$$

$$q_{ij}^{\text{time}} = 2dt \text{Re} \left(c_i^*(t) \langle \chi_i(\mathbf{r}, t) | \partial_t | \chi_j(\mathbf{r}, t) \rangle c_j(t) \right), \quad (8.50)$$

$$q_{ij}^{\text{total}} = \max \{ 0, (q_{ij}^{\text{space}} + q_{ij}^{\text{time}}) \}, \quad (8.51)$$

for transitions of the hidden variable from the j^{th} to the i^{th} eigenstate of $\hat{H}(\mathbf{r}, t)$ during the time interval dt due to non-adiabatic spatial and temporal variations in $\hat{H}(\mathbf{r}, t)$ respectively. This clipping of the individual probabilities is important as

These matrices have zeros along their diagonals, since the above derivation takes into account only probability changes, and does not count probability remaining in the same state as a transition.¹¹ To be able to construct properly stochastic matrices, we can take into account the probability mass that remains in the same state simply by imposing conservation of overall probability, defining a diagonal matrix P^{stay} for the unconditional probabilities of remaining in a state:

$$P_{ii}^{\text{stay}} = |c_i(t)|^2 - \sum_{j \neq i} (P_{ij}^{\text{space}} + P_{ij}^{\text{time}}). \quad (8.52)$$

The sum of all three of these matrices now satisfy the row sum and column sum requirements in order to be the unconditional transition probabilities for a hidden variable theory in the eigenbasis of \hat{H} :

$$P = P^{\text{space}} + P^{\text{time}} + P^{\text{stay}}, \quad (8.53)$$

$$\Rightarrow \sum_j P_{ij} = |c_i(t)|^2, \quad (8.54)$$

$$\sum_i P_{ij} = |c_j(t+dt)|^2. \quad (8.55)$$

¹¹One can see that the $i = j$ term in (8.46) is zero since $|\chi_i\rangle$ is a unit vector, implying its temporal and spatial derivatives must be orthogonal to $|\chi_i\rangle$ itself resulting in a zero inner product.

The corresponding conditional probabilities of a transition to the i^{th} state occurring—given that the hidden variable was already in the j^{th} state—can be obtained via (8.16) as:

$$S_{ij}^{\text{space}} = \frac{1}{|c_j(t)|^2} P_{ij}^{\text{space}}, \quad (8.56)$$

$$S_{ij}^{\text{time}} = \frac{1}{|c_j(t)|^2} P_{ij}^{\text{time}} \quad (8.57)$$

$$S_{ij}^{\text{stay}} = \frac{1}{|c_j(t)|^2} P_{ij}^{\text{stay}}, \quad (8.58)$$

and the sum of these three matrices of conditional probabilities is the overall (left) stochastic matrix for the fewest-switches hidden variable theory:

$$S = S^{\text{space}} + S^{\text{time}} + S^{\text{stay}}. \quad (8.59)$$

In practice we don't need to use this stochastic matrix, though it is instructive to know that the sum of the other conditional probability matrices is indeed a stochastic matrix, such that Tully's fewest-switches is indeed a hidden variable theory. Instead we use the individual matrices in order to distinguish between the different types of transitions that can occur, so that we may make the energy conservation part of the surface hopping algorithm conditional on what type of transition took place.

During a simulation, to choose whether a transition occurs due to spatial or temporal variations in the Hamiltonian, one should not make independent random choices based on the three above matrices of conditional probabilities. Rather one should assemble the probabilities of possible events—transitions from the current state to all others via both spatial and temporal non-adiabatic transitions—into a single list of probabilities, and then take the cumulative sum, resulting in a list of numbers between zero and one. A randomly generated number between zero and one can then be used to determine which event occurs, with the correct probability (Figure 8.3)

Framing fewest-switches as a hidden-variable theory

Tully's fewest-switches algorithm allows one to compute transition probabilities for a hidden variable given the Hamiltonian, the state amplitudes, and a small interval of time. Does this satisfy Aaronson's definition of a hidden variable theory? Not quite, since it requires the Hamiltonian rather than the unitary that describes state vector evolution over a particular time interval, however it is simple to reconcile the two sets of requirements. Given that the interval of time is small, the unitary describing evolution in the local basis can be linked to the effective Hamiltonian in (8.35) via

$$\hat{U}(t + dt, t) = e^{-\frac{i}{\hbar} \hat{H}_{\text{eff}}(r, t) dt}, \quad (8.60)$$

where

$$\hat{H}_{\text{eff}} = \hat{V}(r, t) - i\hbar \hat{U}(r, t) \frac{d}{dt} U_H^\dagger(r, t) \quad (8.61)$$

is the effective Hamiltonian in the eigenbasis of \hat{H} , the matrix representation of which can be extracted from the matrix representation of $\hat{U}(t + dt, t)$ as:

$$H_{\text{eff}}(r, t) dt = i\hbar \text{Log } U(t + dt, t) \quad (8.62)$$

$$\approx i\hbar (\mathbb{I} - U(t + dt, t)) \quad (8.63)$$

where Log is the principal value of the complex matrix logarithm.

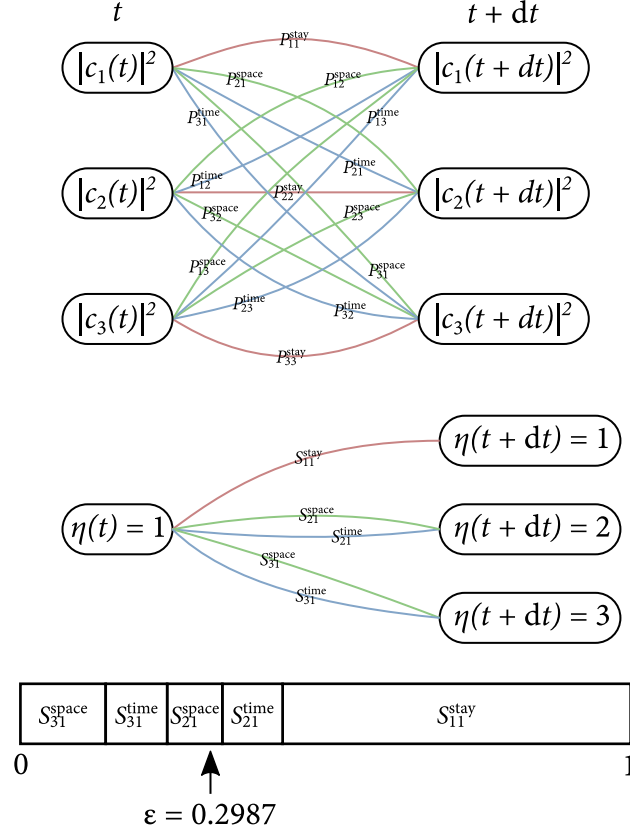


Figure 8.3: Probabilistically choosing a transition. Top: Probability flows between states in a time interval according to the three matrices of unconditional probabilities P^{stay} , P^{space} , and P^{time} in such a way that one total unit of probability is routed from states at the initial time to states at the final time consistent with the populations resulting from quantum mechanical evolution in that timestep. Centre: The hidden variable transitions according to the corresponding conditional probabilities which are elements of the matrices S^{stay} , S^{space} , and S^{time} . Here the hidden variable is in state 1 and we are choosing whether it will remain in that state or if it will transition to state 2 or 3, and if the latter, whether it transitions via a spatial or temporal non-adiabatic transition. Bottom: All the probabilities for what may happen to the hidden variable in a timestep sum to one, so an array of the cumulative probabilities can be constructed, and a random number drawn from the range $[0, 1]$. The index of the smallest element of the array of cumulative sums that the random number is smaller than corresponds to the event to occur. In the above diagrammatic example, the result of the random draw is that the hidden variable is to transition to state 2 via a spatial non-adiabatic transition.

¹²Though we do need its component terms if we wish to distinguish between spatially vs. temporally induced transitions as in [SECREF].

Since only the matrix $H_{\text{eff}}(\mathbf{r}, t) dt$ is required to compute transition probabilities according to (8.43), and not its component terms,¹² specifying the initial state vector and unitary for an interval of time evolution (both in the local basis) is a sufficient input to be able to compute transition probabilities using Tully's fewest-switches algorithm. Writing the resulting matrix of probabilities in terms of U gives:

$$P_{ij} = \begin{cases} \max \left\{ 0, 2 \operatorname{Re} \left(c_i^*(t) U_{ij}(t+dt, t) c_j(t) \right) \right\} & i \neq j \\ |c_i(t)|^2 - \sum_{j \neq i} P_{ij} & i = j \end{cases} \quad (8.64)$$

The corresponding stochastic matrix S can then be obtained by scaling the columns of

P by the state populations as in (8.16). Tully's fewest-switches algorithm thus satisfies Aaronson's definition of a hidden-variable theory provided that the time interval is small. What is somewhat remarkable is the low computational complexity of computing probabilities via fewest-switches. There is no matrix scaling, no matrix permanents, or any other large computational expense. It is not clear how many of Aaronson's axioms are satisfied by fewest-switches, but since it does depend on the actual coupling strengths between states via the non-adiabatic Hamiltonian, one would expect it to satisfy a few of them. Furthermore, the form of fewest-switches when framed in terms of the unitary for the given time interval is remarkably similar to that of Schrödinger theory. In Schrödinger theory one takes the absolute value of elements of U and then applies row and column scalings. In fewest-switches one applies row and column scalings (explicitly given as $c_i(t)^*$ and $c_j(t)$, plus an additional factor of 2), then takes the real part of the result and clips it to zero from below.

8.3.3 Distinguishing between spatial and temporal transitions in arbitrary hidden-variable theories

[TODO flag this as speculative and untested unless I test it]

In the above section [OR SECREP IF I MOVE THIS], we saw how to extract separate stochastic matrices from Tully's fewest-switches algorithm in order to distinguish between non-adiabatic transitions due to motion through space with a spatially varying Hamiltonian as compared to those due to the Hamiltonian's explicit time dependence. This fell naturally out of the derivation simply by not discarding the possible time-dependence of the Hamiltonian and otherwise proceeding similarly to Tully. For an arbitrary hidden variable theory however it is not so simple. In the case of Schrödinger theory for example, we have something of a black box which we feed unitaries and state vectors, and it gives us back stochastic matrices. It is not obvious how one might dig into the details of the matrix scaling to extract two separate conditional probabilities instead of one.

One method might be to compute a unitary as usual for the total evolution of the state vector (excluding decoherence) during a timestep, and compute S^{total} from that according to one's hidden variable theory. Then, repeat this step, but with a unitary computed by holding the Hamiltonian's explicit time dependence constant over the timestep (allowing \mathbf{r} to still vary as normal), to obtain the conditional transition probabilities S^{space} for non-adiabatic transitions due to spatial variations alone.

[I think this won't work due to the possibility of negative probabilities. Probably abandon this and just use Tully's instead. Leave in the paragraph about it not being trivial though! It certainly isn't. Maybe mention that this approach doesn't always work - or maybe you could clip the probabilities to zero. But that would be privileging one type of transition over another.]

8.3.4 Velocity correction and classically disallowed transitions

When a transition of the hidden variable occurs due to a spatial non-adiabatic transition, the kinetic energy of the atom must be adjusted to conserve overall energy. The force on an atom during a transition from the j^{th} state to the i^{th} state due to the spatial non-adiabatic coupling term in the effective Hamiltonian is in the direction of the non-adiabatic coupling vector

$$\mathbf{d}_{ij} = \langle \chi_i(\mathbf{r}, t) | \nabla | \chi_j(\mathbf{r}, t) \rangle \quad (8.65)$$

$$= \left(U_H(\mathbf{r}, t) \nabla U_H^\dagger(\mathbf{r}, t) \right)_{ij} \quad (8.66)$$

where $U_H(\mathbf{r}, t)$ is the matrix representation, in any fixed basis, of the unitary that takes state vectors into the basis in which $\hat{H}(\mathbf{r}, t)$ is diagonal. To conserve energy, the squared component of an atoms velocity in this direction must change by an amount:

$$\Delta v_{ij}^2 = \frac{2}{m} (V_{jj}(\mathbf{r}, t) - V_{ii}(\mathbf{r}, t)), \quad (8.67)$$

where $V_{ii}(\mathbf{r}, t)$ is the adiabatic potential comprising the i^{th} spatially and/or temporally varying eigenvalue of $\hat{H}(\mathbf{r}, t)$. However, it is possible that a change of this size can leave an atom with a negative kinetic energy. Whilst this is quantum-mechanically permissible, it is forbidden classically, and so we simply disallow such transitions, defining a modified matrix of unconditional transition probabilities \tilde{P}^{space} that sets the probability of the disallowed transitions to zero:

$$\tilde{P}_{ij}^{\text{space}} = \begin{cases} P_{ij}^{\text{space}} & \Delta v_{ij}^2 + |\mathbf{v} \cdot \hat{\mathbf{d}}_{ij}|^2 \geq 0 \\ 0 & \Delta v_{ij}^2 + |\mathbf{v} \cdot \hat{\mathbf{d}}_{ij}|^2 < 0 \end{cases} \quad (8.68)$$

where $\hat{\mathbf{d}}_{ij}$ is the unit vector in the direction of \mathbf{d}_{ij} . The diagonal matrix for the probabilities of remaining in each state must be adjusted similarly to absorb the probability discarded this way:

$$\tilde{P}_{ii}^{\text{stay}} = |c_i(t)|^2 - \sum_{j \neq i} (\tilde{P}_{ij}^{\text{space}} + P_{ij}^{\text{time}}) \quad (8.69)$$

The corresponding matrices of conditional probabilities for the hidden variable transitioning, given that it is already in the j^{th} state, are

$$\tilde{S}_{ij}^{\text{space}} = \frac{1}{|c_j(t)|^2} \tilde{P}_{ij}^{\text{space}}, \quad (8.70)$$

$$\tilde{S}_{ij}^{\text{stay}} = \frac{1}{|c_j(t)|^2} \tilde{P}_{ij}^{\text{stay}}, \quad (8.71)$$

Of course, when making probabilistic transitions of the hidden variable, only one column of any of the above matrices is used at a time, so the full matrices \tilde{P}^{space} , \tilde{P}^{time} , \tilde{S}^{space} , and \tilde{S}^{time} do not need to be computed at each timestep.

In the case that a transition of the hidden variable does occur to the i^{th} from the j^{th} state due to a spatial non-adiabatic coupling, the velocity kick to be applied to the classical velocity vector in order to conserve momentum is:

$$\Delta \mathbf{v}_{ij} = \left[\text{sgn}(\mathbf{v} \cdot \hat{\mathbf{d}}_{ij}) \sqrt{(\mathbf{v} \cdot \hat{\mathbf{d}}_{ij})^2 + \frac{2}{m} (V_{ii}(\mathbf{r}, t) - V_{jj}(\mathbf{r}, t))} - (\mathbf{v} \cdot \hat{\mathbf{d}}_{ij}) \right] \hat{\mathbf{d}}_{ij} \quad (8.72)$$

8.4 Decoherence

Was not in Tully's original paper. Prior to discovering the surface hopping literature I came up with simple Markovian decoherence as a mean decoherence of thermal wavelength sized Gaussians [see section]. It is below in [SECREf]. After discovering the surface hopping literature and their fondness for spawning additional trajectories to work out decoherence more accurately, I came up with what I believe is an improvement [see subsection].

In Subsection 8.4.1 I show the general mechanism by which the divergence of trajectories decoheres internal states of an atom, and specifically how this appears as an apparent decrease in the amplitudes of each other state when one imposes the adiabatic trajectory corresponding to one specific internal state.

rev: 122 (f0bd3bb3978d)
author: chrisjbillington
date: Tue Apr 10 17:50:58 2018 +1000
summary: Decoherence stuff

8.4.1 Back action of position measurement on internal state

Consider an atom in a state $|\Psi(t)\rangle$, which is an arbitrary superposition of internal basis states $|\chi_i\rangle$ and motional basis states $|\mathbf{r}\rangle$:

$$|\Psi(t)\rangle = \int \sum_i \psi_i(\mathbf{r}, t) |\chi_i\rangle \otimes |\mathbf{r}\rangle d\mathbf{r}, \quad (8.73)$$

where normalisation requires that

$$\int \sum_i |\psi_i(\mathbf{r}, t)|^2 d\mathbf{r} = 1. \quad (8.74)$$

Recognising that $\psi_i(\mathbf{r}, t)$ is the i^{th} component of a multi-component wavefunction that has one component for each internal state, we define (up to an arbitrary phase factor) a normalised wavefunction $\phi_i(\mathbf{r}, t)$ and corresponding motional state vector $|\phi_i(t)\rangle$ and its coefficient $c_i(t)$ for each component:

$$c_i(t)\phi_i(\mathbf{r}, t) \equiv \psi_i(\mathbf{r}, t) \quad (8.75)$$

$$|\phi_i(t)\rangle \equiv \int \phi_i(\mathbf{r}, t) |\mathbf{r}\rangle d\mathbf{r} \quad (8.76)$$

such that $\int |\phi_i(\mathbf{r}, t)|^2 d\mathbf{r} = 1$, allowing us to write our arbitrary state vector as a sum over the internal basis only:

$$|\Psi(t)\rangle = \sum_i c_i(t) |\chi_i\rangle \otimes |\phi_i(t)\rangle, \quad (8.77)$$

where the spatial state vectors $\{|\phi_i(t)\rangle\}$ are not necessarily orthogonal. To see that spatial separation of the different components leads to decoherence of the internal states, consider the pure density operator corresponding to $|\Psi(t)\rangle$:

$$\hat{\rho}(t) = |\Psi(t)\rangle\langle\Psi(t)| = \sum_{ij} c_i(t)c_j^*(t) |\chi_i(t)\rangle\langle\chi_j(t)| \otimes |\phi_i(t)\rangle\langle\phi_j(t)|, \quad (8.78)$$

which we can write in the $\{|\chi_i\rangle \otimes |\mathbf{r}\rangle\} \equiv \{|\chi_i, \mathbf{r}\rangle\}$ basis, resulting in matrix elements

$$\rho_{ij}(t, \mathbf{r}, \mathbf{r}') = \langle\chi_i, \mathbf{r}|\hat{\rho}(t)|\chi_j, \mathbf{r}'\rangle, \quad (8.79)$$

$$= \psi_i(\mathbf{r}, t)\psi_j^*(\mathbf{r}', t), \quad (8.80)$$

$$= c_i(t)c_j^*(t)\phi_i(\mathbf{r}, t)\phi_j^*(\mathbf{r}', t). \quad (8.81)$$

A partial trace [CITE THE DECOHERENCE BOOK] over the motional degree of freedom results in a reduced density operator describing the measurement statistics of the internal degree of freedom only, assuming ignorance of the position degree of freedom:

$$\hat{\rho}^{\text{red}}(t) = \int \langle\mathbf{r}|\hat{\rho}(t)|\mathbf{r}\rangle d\mathbf{r} \quad (8.82)$$

$$\Rightarrow \rho_{ij}^{\text{red}}(t) = \int \rho_{ij}(t, \mathbf{r}, \mathbf{r}') \delta(\mathbf{r} - \mathbf{r}') d\mathbf{r} \quad (8.83)$$

$$= c_i(t)c_j^*(t) \int \phi_i(\mathbf{r}, t)\phi_j^*(\mathbf{r}, t) d\mathbf{r} \quad (8.84)$$

$$= c_i(t)c_j^*(t) \langle\phi_j(t)|\phi_i(t)\rangle. \quad (8.85)$$

So we see that the off diagonals of the density matrix, representing the coherences of the internal states, are reduced by a factor $\langle\phi_j(t)|\phi_i(t)\rangle$, which will be unity only for pairs

of motional state vectors that are identical. We therefore see that separation in space of the wavefunctions corresponding to different internal states leads to decoherence of the internal states, with decoherence factor $r_{ji} = \langle \phi_j(t) | \phi_i(t) \rangle$. This same decoherence factor appears when—instead of integrating over all positions—we project the total state vector onto a single motional state corresponding to a classical trajectory, which (along with assumptions about the form and evolution of the motional states), is how we impose classicality on the motional degree of freedom in our model. The effect of decoherence when “following” a specific motional state through space in this way is to reduce the amplitudes of all other states not being followed by the decoherence factor between that state and the one being followed, as shown schematically in Figure [FIGREF SCHEMATIC FROM EARLIER SECTION].

Returning to our arbitrary state vector (8.75), we find (ignoring normalisation) the projected state vector $|\tilde{\Psi}(t)\rangle = \hat{R}(t) |\Psi(t)\rangle$, where $\hat{R}(t) = \hat{1} \otimes |\phi_\eta(t)\rangle \langle \phi_\eta(t)|$, resulting from projection of the total state vector $|\Psi(t)\rangle$ onto the specific motional state vector $|\phi_\eta(t)\rangle$ corresponding to the state selected by the hidden variable η . This gives the (non-normalised) state vector one would observe conditional on the particle being in that specific motional state at that specific time.

$$|\tilde{\Psi}(t)\rangle = \hat{R}(t) |\Psi(t)\rangle \quad (8.86)$$

$$= \sum_i \langle \phi_\eta(t) | \phi_i(t) \rangle c_i(t) |\chi_i\rangle \otimes |\phi_\eta(t)\rangle \quad (8.87)$$

$$= \sum_i \tilde{c}_i(t) |\chi_i\rangle \otimes |\phi_\eta(t)\rangle \quad (8.88)$$

$$= |\tilde{\chi}(t)\rangle \otimes |\phi_\eta(t)\rangle, \quad (8.89)$$

where we have defined $\tilde{c}_i(t) = \langle \phi_\eta(t) | \phi_i(t) \rangle c_i(t)$ and $|\tilde{\chi}(t)\rangle = \sum_i \tilde{c}_i(t) |\chi_i\rangle$. We can immediately see that the effect of this projection is to reduce the amplitude of all other states ($i \neq \eta$) by the decoherence factor $r_{\eta i}(t)$. We now proceed by considering this a projective measurement to have occurred at time t , and evolving the resulting projected state vector to time $t + dt$ before projecting it again:

$$|\tilde{\Psi}(t + dt)\rangle = \hat{R}(t + dt) \sum_i \left[\tilde{c}_i(t) - \frac{i}{\hbar} \sum_j H_{ij}(t) \tilde{c}_j(t) dt \right] |\chi_i\rangle \otimes |\phi_i(t + dt)\rangle, \quad (8.90)$$

where $H_{ij}^{\text{red}}(t) = \langle \chi_i | \hat{H}^{\text{red}}(t) | \chi_j \rangle = \langle \chi_i | \phi_\eta(t) | \hat{H}(t) | \chi_j | \phi_\eta(t) \rangle$ are the elements of the “reduced” Hamiltonian $\hat{H}^{\text{red}}(t)$ dictating the dynamics of the internal state, given the imposed motional state $|\phi_\eta\rangle$ and the total Hamiltonian $\hat{H}(t)$ of the system. Note that because of the previous projection already performed, all motional states $|\phi_i(t)\rangle$ were “reset” to be equal to $|\phi_\eta(t)\rangle$ at time t , and therefore the evolved motional state $|\phi_i(t + dt)\rangle$ (the form of which we have not yet specified) represents the evolution of the motional state corresponding to the i^{th} internal state over an interval dt , starting with the initial condition $|\phi_\eta(t)\rangle$. Accordingly, both motional states will still be approximately equal after this short evolution, allowing us to write

$$\langle \phi_\eta(t + dt) | \phi_i(t + dt) \rangle = \langle \phi_\eta(t) | \phi_i(t) \rangle + \frac{d}{dt} \langle \phi_\eta(t) | \phi_i(t) \rangle dt \quad (8.91)$$

$$= 1 + \frac{dr_{\eta i}(t)}{dt}. \quad (8.92)$$

Using this fact in applying the projection in (8.93) yields a product state once more:

$$|\tilde{\Psi}(t + dt)\rangle = \sum_i \left(1 + \frac{dr_{\eta i}(t)}{dt}\right) \left[\tilde{c}_i(t) - \frac{i}{\hbar} \sum_j H_{ij}(t) \tilde{c}_j(t) dt \right] |\chi_i\rangle \otimes |\phi_{\eta}(t + dt)\rangle \quad (8.93)$$

$$\Rightarrow |\tilde{\chi}(t + dt)\rangle = \sum_i \left(1 + \frac{dr_{\eta i}(t)}{dt}\right) \left[\tilde{c}_i(t) - \frac{i}{\hbar} \sum_j H_{ij}(t) \tilde{c}_j(t) dt \right] |\chi_i\rangle \quad (8.94)$$

which we can solve for $\frac{1}{dt} (|\tilde{\chi}(t + dt)\rangle - |\tilde{\chi}(t)\rangle)$ to obtain a differential equation for $|\tilde{\chi}(t)\rangle$:

$$\frac{d}{dt} |\tilde{\chi}(t)\rangle = \left[-\frac{i}{\hbar} \hat{H}^{\text{red}}(t) - \hat{\Gamma}(t) \right] |\tilde{\chi}(t)\rangle, \quad (8.95)$$

where $\hat{\Gamma}(t)$ is a diagonal operator in the $\{|\chi_i\rangle\}$ basis with diagonals

$$\Gamma_i(t) = -\frac{dr_{\eta i}(t)}{dt} = -\frac{d}{dt} \langle \phi_{\eta}(t) | \phi_i(t) \rangle, \quad (8.96)$$

which are *decoherence rates*, and cause exponential damping of all internal states other than the $|\chi_{\eta}\rangle$ state, with faster damping the faster the motional state corresponding to a given internal state diverges with $|\phi_{\eta}(t)\rangle$.

8.4.2 Reframing as continuous weak measurement of internal state

[Write the set of operators and their probabilities. Talk about the speculative stuff here]

This projection onto a single (albeit still quantum) motional state is the first step toward imposing classicality on the motional degree of freedom of our system, and is similar, but not identical to performing a continuous weak quantum measurement of the position degree of freedom. Similarly to the case of continuous weak measurement, we have a set of (time dependent in our case) projection operators $\{\Omega_i(t)\}$ with

$$\Omega_i(t) = \hat{1} \otimes \int \psi_i(\mathbf{r}, t) |\mathbf{r}\rangle \langle \mathbf{r}| d\mathbf{r}, \quad (8.97)$$

and are choosing from them at each timestep with probability of the i^{th} outcome being $\langle \Psi(t) | \Omega_i^\dagger(t) | \Omega_i(t) \rangle$. This is as if we were performing a measurement with hypothetical measurement device not capable of distinguishing different positions perfectly, but only of determining which of the (not necessarily orthogonal) motional state vectors $\{|\phi_i(t)\rangle\}$ the system is in.

Essentially at each timestep we are performing a weak measurement (a positive operator valued measure or POVM [CITE]) of the position degree of freedom, with a hypothetical measurement device that cannot distinguish positions perfectly, and has measurement outcomes corresponding to the set of motional states $\{|\phi_i\rangle\}$.

[ELABORATE. WRITE THE POVMS AND THEIR PROBABILITIES]

The measurement outcome is not selected randomly at the time the projection is performed, but instead pre-ordained by the current value of the hidden variable η . The fact that the hidden-variable makes stochastic transitions in a manner consistent with the Born rule guarantees that the probability of each measurement outcome is the same as if we had chosen an outcome randomly based on the Born rule itself at the time of each measurement. See Section [SECRET SPECULATIVE ASIDE] for a discussion of why use of a hidden variable is preferable to actually using the Born rule.

Effectively performing this measurement at each timestep of a simulation means that we are assuming that any amplitude leaving the envelope of the currently tracked

motional state will either not come back into it, or will have a minimal effect if it does. Given the six-dimensional size of phase space, in which two separated pieces of amplitude of an atom's wavefunction must find each other if they are to interfere, this is a worthy approximation in many cases, to the extent that until recently [CITE] it was not possible to observe such interference in the case of Stern–Gerlach separation of cold atoms. The problem of getting the wavepackets back together again is called the “Humpty Dumpty problem” and was a serious concern to early atom interferometry researches before the use of Bragg pulses [CITE] allowed experiments to create large momentum differences between wavepackets without having them more slowly traverse the intermediate regions of momentum space where they might accumulate coherence-destroying phase noise.

8.4.3 Form of the decoherence factor and the quantum Zeno effect

Finally we will say something about the motional states $\{|\phi_i(t)\rangle\}$. We will assume that each of them is a Gaussian wavepacket with width equal the thermal wavelength, or a specific coherence length we will define, whichever is smaller (see Section [WRITE SECTION HOW BIG IS A WAVEPACKET?]). Each wavepacket moves, without dispersion, with its mean position subject to the adiabatic potential $V_{ii}(\mathbf{r}, t)$ as described in in [SECREP TIME DEPENDENT FORMULATION OF TULLY]. Thus:

$$\langle \mathbf{r} | \phi_i(t) \rangle = \phi_i(\mathbf{r}, t) = A e^{-\frac{|\mathbf{r} - \mathbf{r}_i(t)|^2}{4\sigma^2} + i \frac{m}{\hbar} \mathbf{v}_i(t) \cdot (\mathbf{r} - \mathbf{r}_i(t))} \quad (8.98)$$

$$\frac{d}{dt} \mathbf{r}_i(t) = \mathbf{v}_i(t) \quad (8.99)$$

$$\frac{d}{dt} \mathbf{v}_i(t) = -\frac{1}{m} \nabla V_{ii}(\mathbf{r}_i, t). \quad (8.100)$$

Where A is a real normalisation constant.¹³ This gives for our decoherence factor:

$$R_{ij}(t) = e^{-\left(\frac{1}{8\sigma^2} |\mathbf{r}_{ij}|^2 + \frac{i}{2} \mathbf{r}_{ij}(t) \cdot \mathbf{k}_{ij}(t) + \frac{\sigma^2}{2} |\mathbf{k}_{ij}(t)|^2\right)}, \quad (8.101)$$

where $\mathbf{r}_{ij}(t) = \mathbf{r}_i(t) - \mathbf{r}_j(t)$ and $\mathbf{k}_{ij}(t) = \frac{m}{\hbar} (\mathbf{v}_i(t) - \mathbf{v}_j(t))$ are the relative mean position and wavevector of the pair of wavepackets. As expected, it is equal to unity when the two wavepackets have identical positions and velocities, and decays to zero for increasing relative position and velocity.

[TODO: Flag in intro that I will summarise the necessary equations later in the algorithm section, and make sure you do that]

But now we arrive at a problem. Because all the motional states are reset to be equal to $|\phi_\eta(t)\rangle$ at each timestep, the relative position and wavevector of the two wavepackets are both zero at all times, and so our decoherence rates are [TODO: rename “naive” decoherence rate?]:

$$\Gamma_{ij}(t) = - \left[\frac{dR_{ij}(t)}{dt} \right]_{\substack{\mathbf{k}_{ij}=0 \\ \mathbf{v}_{ij}=0}} \quad (8.102)$$

$$= - \left[\frac{\partial R_{ij}(t)}{\partial \mathbf{r}_{ij}} \cdot \frac{d\mathbf{r}_{ij}(t)}{dt} + \frac{\partial R_{ij}(t)}{\partial \mathbf{k}_{ij}} \cdot \frac{d\mathbf{k}_{ij}(t)}{dt} \right]_{\substack{\mathbf{k}_{ij}=0 \\ \mathbf{v}_{ij}=0}} \quad (8.103)$$

$$= - \left[2\mathbf{k}_{ij} R_{ij}(t) \cdot \frac{d\mathbf{k}_{ij}(t)}{dt} \right]_{\substack{\mathbf{k}_{ij}=0 \\ \mathbf{v}_{ij}=0}} \quad (8.104)$$

$$= 0. \quad (8.105)$$

Every decoherence rate is zero. What is going on? The answer is the quantum Zeno effect [CITE LOADS OF THINGS], which is the name given to the fact that, in the

¹³The overall phase is determined by the offset $\mathbf{r} - \mathbf{r}_i(t)$ in the second term in the exponent, and is chosen such that $\phi_i(\mathbf{r}, t)$ is real at $\mathbf{r} = \mathbf{r}_i$, which ensures that $\langle \phi_i(t) | \phi_j(t) \rangle$ has a phase depending only on the relative positions of the two wavepackets rather than their distance from some arbitrary origin. Acquiring an additional arbitrary phase at each timestep when we apply our reduction operator \hat{R} would not affect the system dynamics, but is unappealing nonetheless.

limit of infinitely frequent measurements of whether quantum evolution has occurred, the back action of the measurement has the effect of preventing the evolution from occurring at all. One saying is that “A watched quantum pot never boils” [CITE?]. In our case, the fact that we are projecting constantly onto our assumed set of motional states causes the amplitude flows to the other motional states to be exactly zero. The appearance of the quantum Zeno effect ought to be a reminder that the assumption of infinitely frequent projective measurements is unphysical. In our case it certainly is—we have merely imagined a hypothetical measurement device collapsing our position states because we don’t want to simulate them, not because any such measurement device actually exists. Nonetheless the internal states of the atom *do* decohere even in the absence of measurement (and there is eventual measurement when the atoms collide with other atoms or otherwise interact with anything in a position-dependent way), and we wish to model the approximate effect of this.

Note that if the decoherence factor had the form of exponential decay:

$$R_{ij}^{\text{Markovian}}(t) = e^{-\gamma_{ij}t}, \quad (8.106)$$

then the decoherence rate would be the constant γ_{ij} , and not zero. This is the case for Markovian decoherence, which is when the environment has no memory of its past interaction with the system. The memory in our case is that the wavepackets accelerate away from each other, starting from zero relative position and velocity at each timestep—the relative position and velocity comprise a memory of the past interaction, which we are erasing each time we reproject.

The lack of an environmental memory is only ever approximately true, and no decoherence factors in nature have the form of a decaying exponential at all times. At small enough times the overlap between two states can only move away from unity quadratically owing to unitary evolution on account of the Schrödinger equation, guaranteeing an initial time derivative of zero for all physical decoherence factors. However in many systems of interest, the specifics of the interaction with the system are quickly forgotten by the environmental states, and the decoherence factor does approach a decaying exponential. This is the case, for example, for spontaneous photon emission by atoms, which one can consider a measurement effect in which the electromagnetic field is being regularly measured by the environment in the photon number basis. In the various quantum trajectories methods used to simulate atoms in the presence of spontaneous emission, if one assumes that the measurements are projective, one must simply assume that the frequent measurements take place at large enough timescales that the decoherence factor is in the exponential decay regime, provided that this is much smaller than other dynamical timescales, which is true for spontaneous emission. This can be recast as a continuous *weak* measurement, rather than infrequent projective measurements [CITE THE STRAIGHTFORWARD INTRODUCTION], which is more physically realistic than the assumption of projective measurements at somewhat arbitrarily chosen intervals, but results in the same differential equations.

A decoherence factor that has the form of a decaying exponential implies that any chosen measurement interval results in the same fractional reduction in state amplitudes, since the differential equation $\frac{d\tilde{c}_i(t)}{dt} = -\gamma\tilde{c}_i(t)$ has the solution $\tilde{c}(t) = e^{-\gamma t}\tilde{c}(0)$, that is, repeated consideration of only the first part of the decoherence factor ends up tracing out the whole decoherence curve over time.

How can we replicate something like this for our separating wavepackets? I have two approaches.

[move this to the other section on non-Markovian decoherence?]

Firstly we need to generalise our conception of the decoherence rate. We have assumed so far that the motional states are being “reset” after each projective measurement to be equal to $|\phi_\eta(t)\rangle$ at each timestep. What is the reduction in the state amplitudes at each

timestep if the wavepackets do not reset, but are left as arbitrary states for the time being? We know that if we were to perform a projective measurement at time t , the i^{th} state amplitude would be reduced by a factor of $R_{\eta i}(t)$ compared to the original state amplitude $c_i(t)$:

$$\tilde{c}_i(t) = R_{\eta i}(t)c_i(t). \quad (8.107)$$

This implies:

$$\frac{d\tilde{c}_i(t)}{dt} = \frac{dR_{\eta i}(t)}{dt}c_i(t) \quad (8.108)$$

$$= \frac{1}{R_{\eta i}(t)} \frac{dR_{\eta i}(t)}{dt} \tilde{c}_i(t). \quad (8.109)$$

and so we see that a decoherence rate that does not reset the environment states is the (negative of the) logarithmic derivative of the decoherence factor, rather than just its derivative. This case encompasses the earlier case in which $R_{\eta i}(t) = 1$ at all times, in which case the logarithmic and ordinary derivatives are the same.

8.4.4 Approximate Markovian decoherence

A crude way to include decoherence is just to [assuming wavepackets are separating with constant acceleration, move this stuff here?] approximate $R_{\eta i}(t)$ as a decaying exponential with approximately the right decay constant. This is crude because $R_{\eta i}(t)$ does not look much like a decaying exponential [see figure FIGREF]. In the limit of large t , its functional form is e^{-t^4} , not the exponential decay required to treat the decoherence as Markovian at any timescale.

To find an approximate Markovian decoherence rate, we first construct a “time ignorant” decoherence factor $\tilde{R}_{\eta i}(t)$ that answers the question “What is the expected value of $R_{\eta i}(t)$ at all times $t > 0$ if I don’t know how long before $t = 0$ the two wavepackets began separating?” We can then take the derivative of this time-ignorant decoherence factor at $t = 0$ to use as an approximate Markovian decoherence rate $\Gamma_i^{\text{Markovian}}$. Whilst extremely approximate, this method of including decoherence is nonetheless an improvement over the Ehrenfest method (which has no decoherence), and over Tully’s original fewest-hops surface hopping method [CITE], which also lacked any decoherence. Others [CITE HEAPS, LOOK AT REVIEW ARTICLE FOR A LIST MAYBE] have since introduced approximate damping of states in a manner very similar to that presented in this subsection, but I was unaware of these developments when I developed this method.

Proceeding, we define the time-ignorant decoherence factor $\tilde{R}_{ij}(t)$ as the average¹⁴ of all decoherence factors one would obtain if the two wavepackets began separating at some point in time before $t = 0$:

$$\tilde{R}_{ij}(t) = A \int_{-\infty}^0 \langle \phi_i(t-t') | \phi_j(t-t') \rangle dt' \quad (8.110)$$

$$= A \int_{-\infty}^0 R(t-t') dt' \quad (8.111)$$

where A is a normalisation constant such that $\tilde{R}_{ij}(0) = 0$, and where we take that $|\phi_i(0)\rangle = |\phi_j(0)\rangle$ with each thereafter evolving according to the classical motion of their centre of mass as in [EQUATIONS].

[Notation is a bit shoddy - specify that motional states are dependent on a trajectory? How can I fit that in the notation?]

¹⁴Not a quantum expectation value, just a weighted sum.

[Mention earlier that the form of $R(t)$ depends on the separation time and is classical evolution of \mathbf{v} and \mathbf{t} . This will fit right in with the definitions of $d\mathbf{v}/dt$ etc.]

[Make gamma have ij subscripts again with the DE - maybe - think about this]

Our time-ignorant decoherence rate is then the (negative of the) derivative of $\tilde{R}_{ij}(t)$ at $t = 0$:

$$\Gamma_i^{\text{Markovian}} = - \frac{\left[\frac{d}{dt} \int_{-\infty}^0 R(t-t') dt' \right]_{t=0}}{\left[\int_{-\infty}^0 R(t-t') dt' \right]_{t=0}}. \quad (8.112)$$

Moving the derivative inside the integral, noting that $\frac{dR(t-t')}{dt} = \frac{dR(t-t')}{dt-t'}$ and setting $t = 0$, we get:

$$\Gamma_i^{\text{Markovian}} = - \frac{\int_{-\infty}^0 R'(-t') dt'}{\int_{-\infty}^0 R(-t') dt'} \quad (8.113)$$

$$= - \frac{\int_0^{\infty} R'(t') dt'}{\int_0^{\infty} R(t') dt'} \quad (8.114)$$

where R' is the derivative of R with respect to its argument. By the fundamental theorem of calculus the numerator is -1 , since $R(t)$ decreases from unity at $t = 0$ to zero as t goes to infinity, leaving us with:

$$\left(\Gamma_i^{\text{Markovian}} \right)^{-1} = \int_0^{\infty} R(t') dt' \quad (8.115)$$

$$= \int_0^{\infty} e^{-\left(\frac{1}{8\sigma^2} |\mathbf{r}_{ij}|^2 + \frac{i}{2} \mathbf{r}_{ij}(t') \cdot \mathbf{k}_{ij}(t') + \frac{\sigma^2}{2} |\mathbf{k}_{ij}(t')|^2 \right)} dt' \quad (8.116)$$

[FIX \mathbf{r}_{ij} - "rel"? Gah.]

Approximating the motion of the wavepackets as moving away from each other with constant acceleration based on the adiabatic potentials at position \mathbf{r} and time t gives:

$$\mathbf{r}_{ij}(t') = \frac{1}{2} \mathbf{a}_{ij}(\mathbf{r}, t) t'^2 \quad (8.117)$$

and

$$\mathbf{k}_{ij}(t') = \frac{m}{\hbar} \mathbf{a}_{ij}(\mathbf{r}, t) t', \quad (8.118)$$

where the relative acceleration is

$$\mathbf{a}_{ij}(\mathbf{r}, t) = -\frac{1}{m} \left(\nabla V_{ii}(\mathbf{r}, t) - \nabla V_{jj}(\mathbf{r}, t) \right). \quad (8.119)$$

8.4.5 Continuously spawned mean auxiliary trajectories

My fancy cool model which is awesome and rad.

$$\mathbf{r}_{i \neq \eta}(t + dt) = \frac{\sum_j \tilde{P}_{ij} \mathbf{r}_j(t)}{|\mathbf{c}_i(t + dt)|^2}, \quad (8.120)$$

$$\mathbf{v}_{i \neq \eta}(t + dt) = \frac{\sum_j \left((P_{ij}^{\text{time}} + \tilde{P}_{ij}^{\text{stay}}) \mathbf{v}_j(t) + \tilde{P}_{ij}^{\text{space}} (\mathbf{v}_i + \Delta \mathbf{v}_{ij}) \right)}{|\mathbf{c}_i(t + dt)|^2}. \quad (8.121)$$

8.5 Algorithm

8.5.1 Further approximations

say that H is based on centre of mass of the eta component

Specifics of the algorithm, now that everything is defined.

- Schrödinger HVSC: with either Markovian decoherence, or auxilliary trajectories. No time dependence allowed
- Fewest hops HVSC/surface hopping: either with Markovian decoherence or auxilliary trajectories. Time dependence allowed!

8.6 results

Pretty results go here

8.7 Discussion and conclusion

Although the method wasn't original, in rediscovering it I have identified that hidden variable theories and hopping algorithms are in fact the same thing. I am relieved that I can use a hidden variable theory that is numerically faster to compute than Schrodinger theory and wonder what the implications are for Aaronson's ideas, given the reasoning outlined in the intro that hidden variable theories ought to be hard to compute. In addition, my decoherence method totally kicks butt and the time dependence thing is nifty.

Discuss how it would make sense for the systems to behave in the presence of collisions w.r.t collapse of state vectors.

8.7.1 Speculative aside: Monte-Carlo wavefunctions and Stochastic Schrödinger equations

The Stochastic Schrödinger equation [CITE] is very similar to the case where the measurement outcome is chosen at projection time, randomly each time rather than according to a hidden variable. This produces the same eventual transition probabilities, but is not suitable for us. If we moved the atom according to the force experienced by the state

8.8 Approximate Markovian decoherence rate for separating wavepackets

Positional separation of two different internal states of an atom leads to decoherence of those states, with a decoherence factor $R_{ij}(t)$ equal to the overlap of the spatial wavefunctions of the two components in question a time t after they began separating. Approximating both wavepackets as initially overlapping Gaussians of width σ , ignoring dispersion, and assuming they separate with constant relative acceleration a_{ij} , the decoherence factor is

$$R_{ij}(t) = \langle \psi_i(t) | \psi_j(t) \rangle \quad (8.122)$$

$$= C \int_{-\infty}^{\infty} e^{-\frac{x^2}{4\sigma^2}} e^{-\frac{(x-x_{rel})^2}{4\sigma^2}} + ik_{rel}x \, dx, \quad (8.123)$$

rev: 122 (f0bd3bb3978d)
author: chrisjbillington
date: Tue Apr 10 17:50:58 2018 +1000
summary: Decoherence stuff

where

$$x_{\text{rel}}(t) = \frac{1}{2}a_{ij}t^2 \quad (8.124)$$

and

$$k_{\text{rel}}(t) = \frac{m}{\hbar}a_{ij}t \quad (8.125)$$

are the wavepackets' relative¹⁵ position and wavenumber due to acceleration for a time t starting from zero relative velocity, and

$$C^{-1} = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} dx \quad (8.126)$$

is a normalisation constant [TODO CHECK IF NEEDS TO BE SQUARED]. Note that this expression holds for any number of dimensions—relative motion is only along one axis so the integrals in all other directions equal one.

Evaluating the Gaussian integral (8.123) gives the following expression for the decoherence factor $R_{ij}(t)$:

$$R_{ij}(t) = e^{-\left[\frac{1}{8\sigma^2}x_{\text{rel}}^2 + \frac{i}{2}x_{\text{rel}}k_{\text{rel}} + \frac{\sigma^2}{2}k_{\text{rel}}^2\right]}. \quad (8.127)$$

This is a decoherence *factor*; it is the factor by which the (i, j) off-diagonal of the reduced density matrix for the atom's internal state will be reduced at time t . The corresponding decoherence *rate* is given by the logarithmic derivative of (8.127):

$$\Gamma_{ij}(t) = -\frac{1}{R_{ij}(t)} \frac{d}{dt} R_{ij}(t). \quad (8.128)$$

The fact that (8.127) does not describe a constant decoherence rate (i.e., it does not have the functional form of exponential decay) means that the back action on the atom's internal state caused by measurements of its motional state will be different depending on the interval of time between measurements.

For example, the logarithmic derivative of (8.127) approaches zero as t goes to zero. This means that in the limit of infinitely frequent measurements, no decoherence occurs at all in between measurements, and the motional state is reset after each measurement such that the wavepackets never separate at all. This is the quantum Zeno effect, and its appearance in models of open quantum systems is usually treated as a reminder that the assumption of infinitely frequent strong measurements is unphysical [CITE].

Since experimentally we are not measuring atoms' motional states so frequently, we ought to wait until the wavepackets are completely separated before performing a projective measurement. As in quantum optics models of open quantum systems, in which the measurement interval “should be large enough to allow the photons to get away from the atom” [CITE The Quantum Jump Approach and Quantum Trajectories Gerhard C. Hegerfeldt], ours should be large enough for the atomic states to get away from each other.

If at large enough times, a decoherence rate is independent of time, that decoherence is called Markovian at that timescale. A Markovian environment is one that has no memory of the decoherence process—it “forgets” any information caused by past interaction with the system. Even though at short times, all decoherence rates in quantum mechanics tend to zero [CITE], if they become Markovian on a timescale shorter than other timescales of interest, the Markov approximation can be used and a constant decoherence rate used

¹⁵ a_{ij} , x_{rel} and k_{rel} are the acceleration, position, and wavenumber of the j^{th} component with respect to the i^{th} component, that is, $a_{ij} = a_j - a_i$, etc.

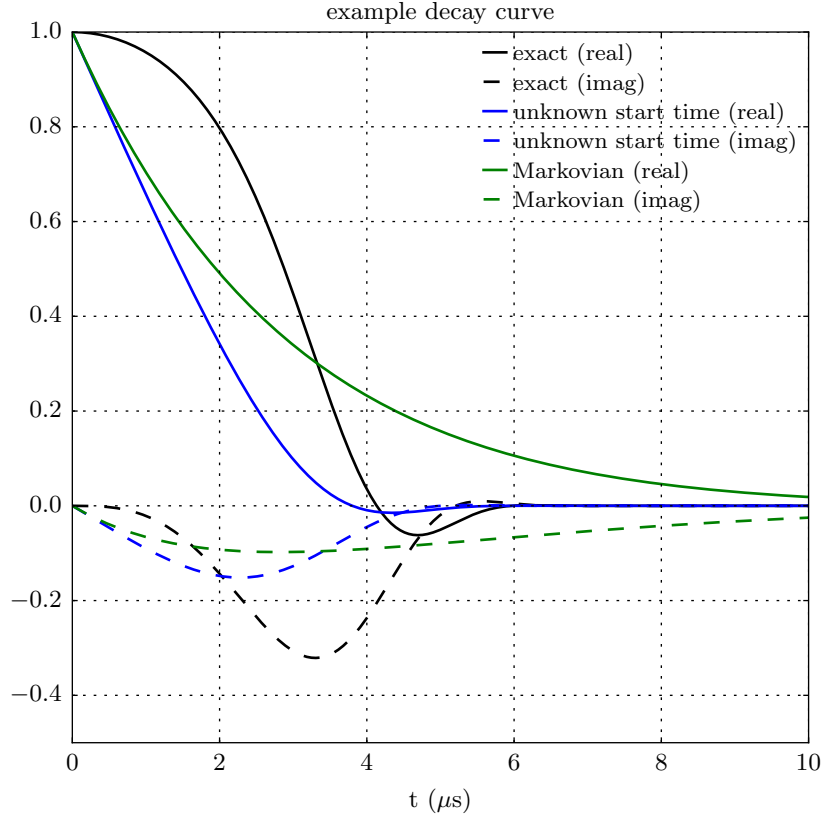


Figure 8.4: Caption.

at all times. In quantum optics, the decoherence factor for the internal state of an atom due to photon emission indeed tends to exponential decay on timescales that are still much shorter than that of the system evolution, and thus the Markov approximation is accurate.

Unlike quantum optics models, our decoherence factor does not describe Markovian decoherence on any timescale. In the limit of large t , its functional form is e^{-t^4} , not the exponential decay required to treat the decoherence as Markovian [CITE] at that timescale. Nonetheless, if we wish to write a time-local differential equation for the internal state of the atom, Markovian decoherence is the only kind we can include [CITE].

To that end, we will now construct a “time ignorant” version of $R_{ij}(t)$ that answers the question “What is the expected decoherence factor at all future times, if you don’t know how long it has been since the two wavepackets began separating?” In this way we can compute an *average* decoherence rate Γ_{ij} described by our decoherence factor, even though $R_{ij}(t)$ does not have a constant decoherence rate at large times. This essentially amounts to finding the best fitting exponential to $R_{ij}(t)$. Whilst this approximation is crude, it is nonetheless an improvement over the Ehrenfest model, which has no decoherence at all (i.e. it has a decoherence rate that is also constant like ours—but equal to zero).

We define the time-ignorant decoherence factor $\tilde{R}_{ij}(t)$ as the overlap of the wavefunction of the i^{th} internal state with a superposition of wavepackets of the j^{th} internal state,

rev: 122 (f0bd3bb3978d)
author: chrisjbillington
date: Tue Apr 10 17:50:58 2018 +1000
summary: Decoherence stuff

with the superposition being over all times in the past the wavepackets began separating:

$$\tilde{R}_{ij}(t) = \langle \psi_i(t) | A \int_{-\infty}^0 |\psi_j(t-t')\rangle dt' \quad (8.129)$$

$$(8.130)$$

where A is a normalisation constant such that $\tilde{R}_{ij}(0) = 1$. Since $|\psi_i(t)\rangle$ is time independent (rather, since we can perform our calculations in the frame of reference in which it is stationary), this is:

$$\tilde{R}_{ij}(t) = A \int_{-\infty}^0 R_{ij}(t-t') dt', \quad (8.131)$$

which is simply the convolution of our decoherence factor with a step function which is nonzero at all negative times. Our average decoherence rate Γ_{ij} is then given by the logarithmic derivative of $\tilde{R}_{ij}(t)$ at $t = 0$:

$$\Gamma_{ij} = -\frac{\tilde{R}'_{ij}(0)}{\tilde{R}_{ij}(0)} \quad (8.132)$$

$$= -\frac{\int_0^\infty \tilde{R}'_{ij}(t) dt}{\int_0^\infty \tilde{R}_{ij}(t) dt} \quad (8.133)$$

$$\Rightarrow \Gamma_{ij}^{-1} = \int_0^\infty e^{-\left[\frac{1}{8\sigma^2}x_{\text{rel}}^2 + \frac{i}{2}x_{\text{rel}}k_{\text{rel}} + \frac{\sigma^2}{2}k_{\text{rel}}^2\right]} dt. \quad (8.134)$$

As mentioned, Γ_{ij} is the decay constant for the best fitting exponential to our decoherence factor (8.127). Although (8.127) looks nothing like a decaying exponential in time, an exponential approximation to it nonetheless ought to decay to zero on the same timescale. An example of this is shown in Figure 8.4

In order to obtain an approximate analytic expression for this integral, we consider two limiting cases and then stitch them together in the intermediate regime. In the limit of small wavepackets, σ is small and thus the first term in the exponent in (8.134) is largest, and the third term is smallest. In this regime, which describes when positional separation (as opposed to separation in k -space) dominates the decoherence, we'll neglect the third term in the exponent and treat the second term as small relative to the first. This gives us:

$$\Gamma_{ij}^{-1}(\text{pos}) \approx \int_0^\infty e^{-\left[\frac{1}{8\sigma^2}x_{\text{rel}}^2 + \frac{i}{2}x_{\text{rel}}k_{\text{rel}}\right]} dt. \quad (8.135)$$

$$\approx \int_0^\infty e^{-\frac{1}{8\sigma^2}x_{\text{rel}}^2} \left(1 - \frac{i}{2}x_{\text{rel}}k_{\text{rel}}\right) dt. \quad (8.136)$$

$$= 2^{\frac{5}{4}}\Gamma(\frac{5}{4})\sqrt{\frac{\sigma}{a_{ij}}} - 2i\frac{m\sigma^2}{\hbar}, \quad (8.137)$$

where we used a first-order Taylor expansion of an exponential in (8.136). We similarly use a first order expansion to take the reciprocal of (8.137) (since the second term is much smaller than the first¹⁶), and arrive at:

$$\Gamma_{ij}(\text{pos}) \approx \frac{1}{2^{\frac{5}{4}}\Gamma(\frac{5}{4})}\sqrt{\frac{a_{ij}}{\sigma}} + \frac{i}{2\sqrt{2}\Gamma(\frac{5}{4})^2}\frac{m\sigma a_{ij}}{\hbar} \quad (8.138)$$

Similarly for the large σ regime, we neglect the first term in the exponent of (8.134) and consider the second term small relative to the third. This is the regime in which the

¹⁶This isn't necessary in order to obtain a simple expression for $\Gamma_{ij}(\text{pos})$ —the reciprocal without this approximation is equally simple—but it leaves us with power laws for the real and imaginary parts of $\Gamma_{ij}(\text{pos})$, which are easier to stitch together with those from the large σ regime.

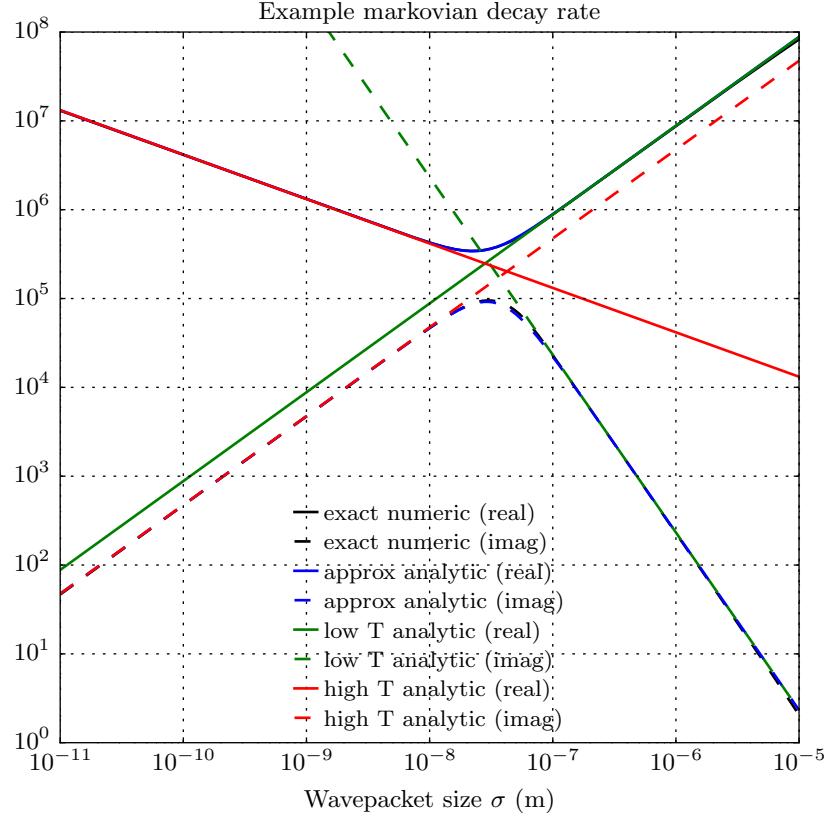


Figure 8.5: Caption.

decrease in overlap of the two wavepackets is dominated by their separation in velocity space. Following the same process as above gives:

$$\Gamma_{ij}^{-1}(\text{vel}) \approx \int_0^\infty e^{-\left[\frac{i}{2}x_{\text{rel}}k_{\text{rel}} + \frac{\sigma^2}{2}k_{\text{rel}}^2\right]} dt. \quad (8.139)$$

$$\approx \int_0^\infty \left(1 - \frac{i}{2}x_{\text{rel}}k_{\text{rel}}\right) e^{-\frac{\sigma^2}{2}k_{\text{rel}}^2} dt \quad (8.140)$$

$$= \sqrt{\frac{\pi}{2}} \frac{\hbar}{m\sigma a_{ij}} - i \frac{\hbar^3}{2m^3\sigma^4 a_{ij}^2} \quad (8.141)$$

$$\Rightarrow \Gamma_{ij}(\text{vel}) \approx \sqrt{\frac{2}{\pi}} \frac{m\sigma a_{ij}}{\hbar} + \frac{i}{\pi} \frac{\hbar}{m\sigma^2} \quad (8.142)$$

Equations (8.138) and (8.142) are our final expressions for the decoherence rate in the limit of small and large wavepackets respectively. Adding their real parts in quadrature and adding the reciprocals of their imaginary parts then provides a reasonable approximation for Γ_{ij} over all wavepacket sizes:

$$\Gamma_{ij} \approx \left[\text{Re}(\Gamma_{ij}(\text{pos}))^2 + \text{Re}(\Gamma_{ij}(\text{vel}))^2 \right]^{\frac{1}{2}} + i \left[\text{Im}(\Gamma_{ij}(\text{pos}))^{-1} + \text{Im}(\Gamma_{ij}(\text{vel}))^{-1} \right]^{-1}. \quad (8.143)$$

We now have an approximate analytic expression that is computationally inexpensive to evaluate for each atom in an ensemble at every timestep of a differential equation. An

example showing the accuracy of (8.143), compared to the exact expression (8.134) for Γ_{ij} over a range of wavepacket sizes is shown in Figure 8.5.

8.9 Choice of wavepacket size

How big is a wavepacket? Be sure to say it's the thermal wavelength except not too big because the single mode approximation won't be valid otherwise.

This page intentionally left blank

References

rev: 122 (f0bd3bb3978d)
author: chrisjbillington
date: Tue Apr 10 17:50:58 2018 +1000
summary: Decoherence stuff

This page intentionally left blank

Word count

Total

Words in text: 46935

Words in headers: 412

Words outside text (captions, etc.): 6801

Number of headers: 96

Number of floats/tables/figures: 34

Number of math inlines: 1787

Number of math displayed: 303

Files: 9

Subcounts:

text+headers+captions (#headers/#floats/#inlines/#displayed)
7422+77+1300 (23/3/559/73) File(s) total: atomic_physics.tex
725+17+385 (4/4/7/0) File(s) total: experiment.tex
50+0+0 (0/0/0/0) File(s) total: front_matter.tex
10014+117+834 (24/5/306/89) File(s) total: hidden_variables.tex
1596+3+98 (2/0/12/0) File(s) total: introduction.tex
23087+155+3403 (31/15/824/137) File(s) total: numerics.tex
4+6+0 (1/0/0/0) File(s) total: software.tex
4001+24+781 (8/7/79/4) File(s) total: velocimetry.tex
36+13+0 (3/0/0/0) File(s) total: wave_mixing.tex

rev: 122 (f0bd3bb3978d)
author: chrisjbillington
date: Tue Apr 10 17:50:58 2018 +1000
summary: Decoherence stuff