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 Helmerson
Dr Lincoln Turner
Dr Russell Anderson

School of Physics and Astronomy
Monash University

April, 2018

This page intentionally left blank

Contents

This page intentionally left blank

Hidden variables for semiclassical models with state-dependent forces

Hidden-variable theories [1, 2] are interpretations of quantum mechanics that posit definite states underlying quantum state vectors, such that quantum indeterminacy is an illusion—an emergent phenomenon rather than a fundamental fact. Bell's theorem [3] 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 [4] 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-variable 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 can be inaccurate—namely the Stern–Gerlach experiment and similar situations in which considerable entanglement between motional and internal degrees of freedom of atoms can develop.

In Section 7.1 I give an overview of what a semiclassical method is, their most common implementation and in what situations this is insufficiently accurate, motivating the need for an improved method. In Section 7.2 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 7.3 I then go into the implications of welcoming a hidden variable into a semiclassical model, including additional required assumptions and approximations, and in Sections 7.4 and 7.5 I go into detail deriving the equations of motion for the model and presenting some algorithmic and computational details. Finally in Section 7.6, having provided all the background arguments and details, I present the complete algorithm(s), before showing simulation results in Section 7.7 that compare the model to the underlying exact Schrödinger wave
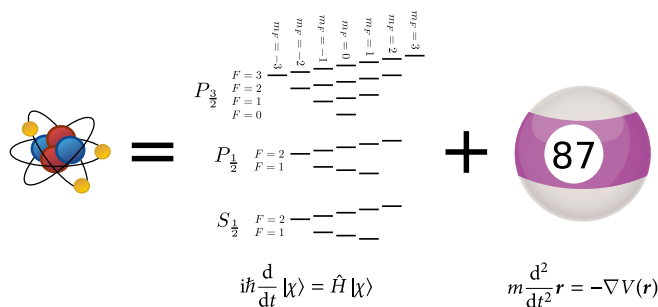
Figure 7.1: Artist's depiction of a semiclassical atom.

equation, and concluding in Section 7.8 with further discussion of the method's benefits and limitations.

During writing this thesis, I discovered that the ideas underpinning this method are (perhaps unfortunately) not original, and that the 'surface hopping' method enjoys widespread use and continued development in the field of computational chemical physics [5–15]. A preprint of an early version of my method [16] bears a striking resemblance to a paper by Tully [5], the pioneer of surface hopping methods. Use of these ideas do not however appear to be common in cold atom physics, but with direct simulation of evaporative cooling and laser cooling schemes becoming increasingly plausible given the increasing availability of computational power as well as efficient molecular dynamics techniques such as Direct Simulation Monte Carlo (DSMC) [17], this sort of convergent evolution of numerical techniques is perhaps not surprising. Throughout this chapter I make comparisons between my own methods and those in the existing surface hopping literature. Some long-standing limitations of my method may be resolved in light of what I have now read in the surface hopping literature, leading to speculative improvements that I have yet to test. These improvements are outlined in this chapter, and flagged as speculative.

7.1 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 of atoms 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 the centre of mass motion of the (relatively heavy) atom, but not for the relative motion of its (much lighter) electrons with respect to the nucleus, or for the nuclear and electronic spin degrees of freedom [5].

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 $\boldsymbol{r}$ and velocity $\boldsymbol{v}$ evolving according to Newton's second law (Figure 7.1).

Once one has defined a potential function $V(\boldsymbol{r})$ and a Hamiltonian (also possibly

varying with space) $\hat{H}(\boldsymbol{r})$, and other possible additions,[1] 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 $\boldsymbol{r}$ and $\boldsymbol{v}$ for motional degree of freedom in time according to the coupled differential equations

$$\frac{\mathrm{d}}{\mathrm{d}t}|\chi\rangle = -\frac{\mathrm{i}}{\hbar}\hat{H}(\boldsymbol{r})|\chi\rangle, \tag{7.1}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{v} = -\frac{1}{m}\nabla V(\boldsymbol{r}), \tag{7.2}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{r} = \boldsymbol{v}. \tag{7.3}$$

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

7.1.1 Stern–Gerlach separation and evaporative cooling

In the Stern–Gerlach experiment [20], 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 (if the atoms are spin-$\frac{1}{2}$) of positions 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 outcome of the Stern–Gerlach experiment is a consequence of quantum mechanics, to be sure, but it has little to do with the wave nature of the atoms themselves. 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 confined 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. 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 $F \propto -\nabla|\boldsymbol{B}|$ 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

[1]Such as using a Monte-Carlo wavefunction method [18, 19] to model the effect of spontaneous emission on $|\chi\rangle$, and modifying $\boldsymbol{v}$ instantaneously by a random-direction recoil velocity upon each photon emission.

local field and remain spin-down, even as the field as seen by each atom fully reverses its direction every half orbital period.

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 [21, 22]) 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 (7.1) to (7.3) are not sufficient, because exists no single classical potential $V(\boldsymbol{r})$ 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: $V = V(\boldsymbol{r}, |\chi\rangle)$. Ehrenfest's theorem [23] states that

$$m \frac{\mathrm{d}^2}{\mathrm{d}t^2} \langle \hat{\boldsymbol{r}} \rangle = -\langle \nabla \hat{V} \rangle, \tag{7.4}$$

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

$$m \frac{\mathrm{d}^2}{\mathrm{d}t^2} \boldsymbol{r} = -\nabla \langle \chi | \hat{V}(\boldsymbol{r}) | \chi \rangle, \tag{7.5}$$

where the operator $\hat{V}(\boldsymbol{r})$ 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

$$m \frac{\mathrm{d}^2}{\mathrm{d}t^2} \boldsymbol{r} = -\nabla \langle \chi | \hat{H}(\boldsymbol{r}) | \chi \rangle, \tag{7.6}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}t} |\chi\rangle = -\frac{\mathrm{i}}{\hbar} \hat{H}(\boldsymbol{r}) |\chi\rangle, \tag{7.7}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{v} = -\frac{1}{m} \langle \chi | \hat{H}(\boldsymbol{r}) | \chi \rangle, \tag{7.8}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{r} = \boldsymbol{v}. \tag{7.9}$$

The Ehrenfest semiclassical method is the same as the simple semiclassical method (7.1) to (7.3), except that it has an answer to the question "What should we use for $V(\mathbf{r})$ when 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 or a Majorana spin flip in a magnetic trap, the expectation value of position is a poor match to reality. In the Stern–Gerlach experiment beginning with spin-polarised atoms, a trajectory subject to the mean force or potential (which are both zero for a 50 : 50 superposition) would land in a single blob in the middle of the screen, rather than two blobs displaced from the centre. In the case of an atom approaching the field zero in a magnetic trap, use of the mean force would result in the atom broadening its orbit somewhat, rather than splitting into multiple possible trajectories (Figure 7.2). Semiclassical simulations of evaporative cooling performed by Christopher Watkins (unpublished) 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.

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 any parameter, variable or label specifying which state has been chosen is a *hidden variable*.

7.2 Hidden-variable theories

In his paper *Quantum computing and dynamical quantum mdels*, Aaronson defines a *dynamical quantum model* [24]:

> 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 to *hidden-variable theories* and instead defines [2]:
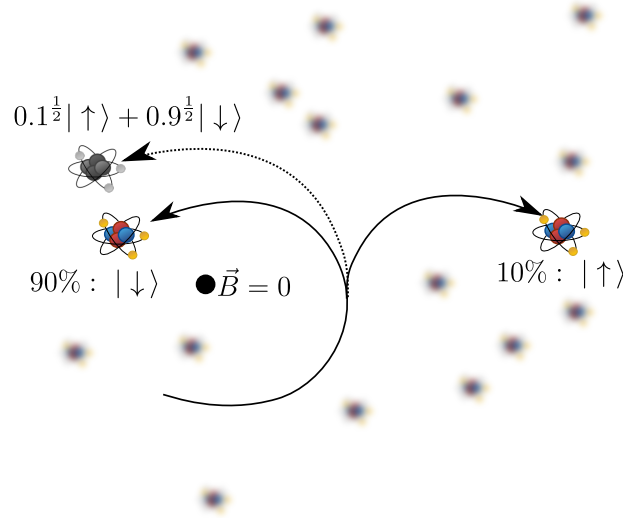
Figure 7.2: When atoms pass near to the field zero in a magnetic trap, their wavepackets diverge in space as multiple trajectories, each corresponding to one local spin projection state. For example, if a spin-$\frac{1}{2}$ atom undergoes a partial spin-flip such that it has a ten percent probability of being spin-up, the end result will be two approximate trajectories. One trajectory describes the motion of a wavepacket that is fully spin-up, and one fully spin-down, with the spin-down wavepacket's squared amplitude equal to 0.1. The Ehrenfest semiclassical method however can only model a single trajectory, and the end result when using this method is a single trajectory being approximately the mean of the two actual trajectories, and retaining a 90 : 10 ratio of spin-down:spin-up state populations.

> 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 first paper is superseded by the second, 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), probabilistically depending on the change in quantum populations of the 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 $\{|\chi_i\rangle\}$ according to a unitary $\hat{U}(t', t)$:

$$
\begin{bmatrix} c_1(t') \\ c_2(t') \\ c_3(t') \\ \vdots \end{bmatrix} = \begin{bmatrix} U_{11}(t',t) & U_{12}(t',t) & U_{13}(t',t) & \dots \\ U_{21}(t',t) & U_{22}(t',t) & U_{23}(t',t) & \dots \\ U_{31}(t',t) & U_{32}(t',t) & U_{33}(t',t) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} c_1(t) \\ c_2(t) \\ c_3(t) \\ \vdots \end{bmatrix}, \tag{7.10}
$$

where $c_i = \langle \chi_i | \chi \rangle$ and $U_{ij}(t', t) = \langle \chi_i | \hat{U}(t', t) | \chi_j \rangle$, 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 $|\chi(t)\rangle$ and unitary $\hat{U}(t', t)$ in the $\{ |\chi_i\rangle \}$ basis, that satisfies

$$
\begin{bmatrix} |c_1(t')|^2 \\ |c_2(t')|^2 \\ |c_3(t')|^2 \\ \vdots \end{bmatrix} = \begin{bmatrix} S_{11}(t', t) & S_{12}(t', t) & S_{13}(t', t) & ... \\ S_{21}(t', t) & S_{22}(t', t) & S_{23}(t', t) & ... \\ S_{31}(t', t) & S_{32}(t', t) & S_{33}(t', t) & ... \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} |c_1(t)|^2 \\ |c_2(t)|^2 \\ |c_2(t)|^2 \\ \vdots \end{bmatrix}. \tag{7.11}
$$

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 $|\chi_\eta\rangle \in \{ |\chi_i\rangle \}$ assigned by the hidden variable will change from one to another in the given time interval:

$$
\Pr(\eta(t')\mathord{=}i|\eta(t)\mathord{=}j) = S_{ij}(U(t', t), \chi(t)). \tag{7.12}
$$

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

$$
\Pr(\eta(t')\mathord{=}i|\eta(t)\mathord{=}j) = |c_i(t')|^2, \tag{7.13}
$$

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 jumps between states randomly based on their population, with no regard for its history or whether there were actually amplitude flows between the states between which it is transitioning. Nonetheless, this theory, called *product theory* [2], 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 some additional axioms [2] 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 (7.12) 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:

$$
\Pr(\eta(t')\mathord{=}i) = \sum_j \Pr(\eta(t')\mathord{=}i|\eta(t)\mathord{=}j) \Pr(\eta(t)\mathord{=}j), \tag{7.14}
$$

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

$$
|c_i(t')|^2 = \sum_j S_{ij}(U(t', t), \chi(t))|c_j(t)|^2. \tag{7.15}
$$

We now define the matrix of absolute transition probabilities P:

$$
P_{ij} = S_{ij}(U(t', t), \chi(t))|c_j(t)|^2, \tag{7.16}
$$

such that:

$$
|c_i(t')|^2 = \sum_j P_{ij}. \tag{7.17}
$$

So we have that P has row sums equal to the final squared amplitudes. And, because $P_{ij} = S_{ij}|c_j(t)|^2$ and S is a stochastic matrix with column sums equal to one,[2] 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 [2]. The idea is to form P by starting with the matrix of absolute values of U, and simply scaling its rows and columns to have the correct values:[3]

$$P = \begin{bmatrix} a_1 b_1 |U_{11}(t',t)| & a_1 b_2 |U_{12}(t',t)| & a_1 b_3 |U_{13}(t',t)| & \dots \\ a_2 b_1 |U_{21}(t',t)| & a_2 b_2 |U_{22}(t',t)| & a_2 b_3 |U_{23}(t',t)| & \dots \\ a_3 b_1 |U_{31}(t',t)| & a_3 b_2 |U_{32}(t',t)| & a_3 b_3 |U_{33}(t',t)| & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{7.18}$$

that is,

$$P_{ij} = a_i b_j |U_{ij}(t',t)|, \tag{7.19}$$

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

$$\sum_j a_i b_j |U_{ij}(t',t)| = |c_i(t')|^2, \tag{7.20}$$

$$\sum_i a_i b_j |U_{ij}(t',t)| = |c_j(t)|^2, \tag{7.21}$$

which can be solved numerically, and then the Schrodinger theory stochastic matrix then able to be extracted by inverting (7.16):

$$S_{ij}(U(t',t), \chi(t)) = a_i b_j \frac{|U_{ij}(t',t)|}{|c_j(t)|^2}. \tag{7.22}$$

Schrödinger theory is the hidden-variable theory I have used in the simulations in the results section of this chapter (Section 7.7). In Section 7.4.1 I detail my efforts to most efficiently find the row and column scalings in order to actually compute the Schrödinger theory S matrix.

Having now discovered Tully's fewest-switches algorithm in the surface hopping literature [5, 7, 12] (also detailed in Section 7.4.2), I no longer recommend using Schrödinger theory as the hidden-variable theory in a hidden-variable semiclassical/surface hopping simulation. Tully's fewest-switches is superior in two aspects. As detailed in Section 7.4.1, Schrödinger theory is computationally expensive to evaluate (except in the case of a two-state system), and does not easily allow a distinction to be drawn between motion of a particle through an inhomogeneous field as the cause of non-adiabatic transitions, versus non-adiabatic transitions being caused by explicit time dependence of said fields. The usefulness of this distinction will hopefully be made clear in Sections 7.4.2 and 7.4.3, though since I have not been able to make the distinction in the past using only Schrödinger theory, these improvements are untested and hence speculative.

7.3   Overview of method

To motivate the method and highlight some necessary properties that any sensible semi-classical model capable of simulating Stern–Gerlach separation must have, consider, as an example, a spin-$\frac{1}{2}$ particle undergoing Stern–Gerlach separation in a magnetic field gradient. The two spin components accelerate away from each other, with the relative acceleration vector pointing in the same direction as the gradient in field strength. We can

[2]In 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.

[3]One 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 theory to agree with quantum probabilities.
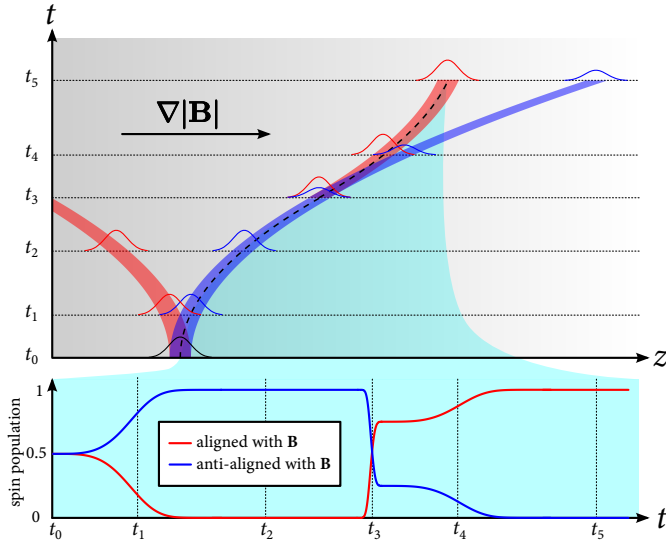
Figure 7.3: Schematic of method. At t_0 an atom is in a 50 : 50 superposition of spin-up:spin-down population, and the two spin components accelerate apart in the magnetic field gradient. If the hidden variable dictates that we follow the spin-down component (the initial trajectory given by the dotted line), then we see a reduced spin-up population at times t_1 and t_2. At t_3 the field changes direction suddenly, and partly flips the spins (in the local basis with quantisation axis given by the direction of the magnetic field). The stochastic hidden variable transitions to instead select the spin-up component, which we follow thereafter. Following the spin-up component we then see the a reduced spin-down population at t_5 due to the wavepackets separating once more.

ask the question: "What spin-state populations would I see, if I followed the spin-down wavepacket only?" The meaning of "follow" will be made more precise in Section 7.5, but for now we will simply look at the spin populations in the vicinity of the region of space occupied by the spin-down wavepacket.

As shown in Figure 7.3, one sees the spin-up population decreasing over time until only spin-down population remains. The rate at which spin-up population 'leaves' the region of space we are watching depends on how quickly the wavepackets are accelerating away from each other, as well as the shape of the wavepackets. Likewise, if one follows the spin-up component instead, one sees the spin-down population decrease to zero.

Motivated by this observation, the hidden-variable semiclassical method is a phenomenological model comprised of the following bolted-together pieces:

- One internal state of the atom is chosen at any moment in time, stored along with the state vector as a stochastic hidden variable that makes transitions at each timestep, according to a hidden-variable theory. The hidden-variable theory takes as input the state vector and Hamiltonian or unitary evolution matrix describing the evolution of the internal state of the atom at each timestep, and thus the probability of the hidden variable selecting out a particular internal state is equal to that state's population at each moment in time.

- The internal state of the atom evolves according to the Schrödinger equation, but with the addition of back-action caused by the continuous projection of the atomic wavefunction onto one motional state: a thermal-wavelength sized Gaussian wavepacket with centre of mass motion evolving classically according to the adiabatic potential experienced by the internal state selected by the hidden variable.

- Upon a change in the state selected by the hidden variable, the velocity of the atom is modified instantaneously as required by energy conservation, or if this would result in a negative kinetic energy, the transition is disallowed.

Written in the same way as we previously described semiclassical models, the equations of motion for the model are the following coupled differential equations and stochastic transition rule:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,|\tilde{\chi}\rangle = -\frac{\mathrm{i}}{\hbar}\hat{H}(\boldsymbol{r},t)\,|\tilde{\chi}\rangle - \hat{\Gamma}_{\eta}(\boldsymbol{r},t)\,|\tilde{\chi}\rangle, \tag{7.23}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{v} = -\frac{1}{m}\nabla V_{\eta}(\boldsymbol{r},t), \tag{7.24}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{r} = \boldsymbol{v} \tag{7.25}$$

$$\Pr(\eta(t')\!=\!i|\eta(t)\!=\!j) = S_{ij}(U_{\mathrm{eff}}(t',t),\chi(t)), \tag{7.26}$$

where $|\tilde{\chi}\rangle$ is the (non-normalised) internal state vector including the effect of back-action, $\hat{\Gamma}_{\eta}(\boldsymbol{r},t)$ is a non-Hermitian operator that implements this back-action by decaying the amplitude of states not being followed (there are many ways to do this, all approximate, discussed in Section 7.5), η is the hidden variable, V_{η} is the adiabatic potential experienced by the eigenstate selected by the hidden variable, χ is the vector representation of the (normalised) state vector in the local eigenbasis, and $U_{\mathrm{eff}}(t',t)$ is the unitary matrix describing the evolution of the state vector from time t to t' in the local eigenbasis under the action of the effective Hamiltonian $\hat{H}_{\mathrm{eff}}$, defined in Section 7.4.2. In addition to these evolution rules, the state vector must be normalised at each timestep of simulation,[4] and the velocity of the atom modified instantaneously whenever a transition is made in order to conserve energy. This latter requirement is detailed in Section 7.4.3.

[4] It would be relatively simple to include a term in the differential equation to preserve normalisation, but the non-normalisation-preserving differential equation is simpler to write, compute, and understand.

7.4   Hidden variables: implementation details

In this section I go into the gritty details of numerically evaluating hidden-variable theories, conserving energy when a transition occurs, and some other concerns.

7.4.1   Numerically evaluating Schrödinger theory

As mentioned, I now consider Schrödinger theory a worse choice, compared to Tully's fewest-switches, for a hidden-variable theory to use in practice for a hidden-variable semiclassical model, primarily due to the high computational cost of evaluating it. Nonetheless here I present the results of my investigation into how to minimise this computational cost.

There are many ways to solve numerically for the row and column scalings required to compute the S matrix of Schrödinger theory (7.22), 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 (7.20) for all a_i, and solving (7.21) for all b_i, until the result converges. This is called the Sinkhorn–Knopp method of r-c (row-column) scaling [25], but is computationally intensive, with slow convergence [26]. An alternative is the method by Linial et al. [26] 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 a considerable speedup.

[5] The 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_ib_j appear in the resulting scaled matrix

[6] With 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 $\{b_i\}$ by a constant may be necessary to prevent the values numerically overflowing or underflowing in the middle of a simulation.

For the case of a two-state system, the row and column scalings can be found analytically, with the result

$$a_1 = 1 \tag{7.27}$$

$$a_2 : a_2^2 + \left(\frac{1}{AB} - \frac{A}{B}|c_1(t)|^2 - \frac{B}{A}|c_2(t)|^2 \right) a_2 + \frac{|c_2(t')|^2}{|c_1(t')|^2} = 0 \tag{7.28}$$

$$b_1 = \frac{|c_1(t')|^2}{A + Ba_2} \tag{7.29}$$

$$b_2 = \frac{|c_2(t')|^2}{B + Aa_2}, \tag{7.30}$$

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 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 two-state 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 $|U_{ij}|$, $|c_i(t')|^2$ and $|c_j(t)|^2$ 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 (7.20) and (7.21) or minimising the residuals [27]. 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 that 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 reference [27], one could construct a hybrid method that takes a Newton step, and then checks the row sum and column sum residuals, and if they increased 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.

A final note is that my hidden-variable 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 [16] was limited to two-component systems, and the probability of transition was computed as:

$$\Pr(\eta(t')\!=\!2|\eta(t)\!=\!1) = \max\left(0, |c_2(t')|^2 - |c_2(t)|^2\right), \tag{7.31}$$

$$\Pr(\eta(t')\!=\!1|\eta(t)\!=\!2) = \max\left(0, |c_1(t')|^2 - |c_1(t)|^2\right), \tag{7.32}$$

[7] In 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 $20 - 40\,\mu s$ per matrix on my computer. This is when transitions are actually occurring; the algorithm converges in zero or one step when the unitary is the identity and the state vector is in a single eigenstate, before and after periods of non-adiabatic evolution.

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 the most obvious of Aaronson's axioms so as not to behave like the product theory. I chose Schrödinger theory fairly arbitrarily, it being the one that seemed most easily computable out of the two presented in Aaronson's paper [2], but one might try using Aaronson's flow theory, or inventing another altogether. The surface hopping literature uses Tully's fewest-switches algorithm with much success, as well as approximations to it [29] that are even cheaper, computationally speaking.

Interestingly, the main conclusion of Aaronson's paper [2] 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 computer. It is therefore perhaps not surprising that hidden-variable theories ought to be computationally expensive to simulate on a classical computer. Prior to discovering Tully's fewest-switches algorithm, I was therefore somewhat resigned to the fact that any hidden-variable theory would likely be computationally expensive to compute. In light of this it was pleasantly surprising to discover that Tully's fewest-switches (discussed in the next subsection) is computationally cheap. The seeming conflict could be reconciled however if the computational power of hidden variables is crucially dependent on some feature we are not interested in and which fewest-switches does not capture, such as agreement with arbitrary unitaries rather than only those due to evolution over small time intervals, as is assumed by fewest-switches.

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

In this subsection I derive Tully's fewest-switches algorithm [5, 7] 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 atom 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 atom by the driving field without conserving its total energy. Thus, when a time dependent potential is used in a hidden-variable semiclassical simulation, I propose that velocity corrections ought to only be performed following transition of the hidden variable if that transition was due to spatial variation in the Hamiltonian, and not temporal variation.

As flagged earlier, this proposed improvement is speculative and untested. In the results section of this chapter I have only tested the model for the case of time-independent fields. Though accepting this proposed improvement would clearly yield the correct behaviour in both the limit of a time-independent inhomogeneous field (energy is always conserved via velocity jumps) and a time-dependent homogeneous field (energy is never required to be conserved and there are no velocity jumps); for a field with both time and space dependence it is not clear that giving some atoms velocity jumps and some not is particularly accurate, though it is certainly an improvement over giving velocity jumps to them all. Perhaps upon comparison with the Schrödinger wave equation it will become clear that say, giving all atoms velocity jumps of a smaller magnitude yields good results. How much smaller, however, will likely depend on the same matrix elements computed below. This subsection also serves to present Tully's fewest-switches algorithm and the concepts and notation related to it that I refer to in later sections.

[8] This was before I coincidentally came across the definition of a stochastic hidden-variable theory in Aaronson's book *Quantum computing since Democritus* [28] and realised that what I had made was a hidden-variable theory, allowing me to choose a more general one from his paper (and before later still, finding Tully's fewest-switches algorithm).

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

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\chi(t)\rangle = \hat{H}(\boldsymbol{r},t)|\chi(t)\rangle \tag{7.33}$$

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

$$\hat{H}(\boldsymbol{r},t) = \hat{U}_H^\dagger(\boldsymbol{r},t)\hat{V}(\boldsymbol{r},t)\hat{U}_H(\boldsymbol{r},t), \tag{7.34}$$

where $\hat{V}(\boldsymbol{r},t)$ 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 $\hat{H}$ as[9]

$$|\chi_H(t)\rangle = \hat{U}_H(\boldsymbol{r},t)|\chi(t)\rangle \tag{7.35}$$

$$\Rightarrow |\chi(t)\rangle = \hat{U}_H^\dagger(\boldsymbol{r},t)|\chi_H(t)\rangle. \tag{7.36}$$

Substituting (7.35) into (7.33) and premultiplying by $\hat{U}_H(\boldsymbol{r},t)$ yields

$$i\hbar\hat{U}_H(\boldsymbol{r},t)\frac{\mathrm{d}}{\mathrm{d}t}\hat{U}_H^\dagger(\boldsymbol{r},t)|\chi_H(t)\rangle = \hat{U}_H(\boldsymbol{r},t)\hat{H}(\boldsymbol{r},t)\hat{U}_H^\dagger(\boldsymbol{r},t)|\chi_H(t)\rangle, \tag{7.37}$$

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

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\chi_H(t)\rangle = \left[\hat{V}(\boldsymbol{r},t) - i\hbar\hat{U}_H(\boldsymbol{r},t)\frac{\mathrm{d}}{\mathrm{d}t}U_H^\dagger(\boldsymbol{r},t)\right]|\chi_H(t)\rangle \tag{7.38}$$

$$= \hat{H}_{\text{eff}}|\chi_H(t)\rangle. \tag{7.39}$$

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 the *adiabatic basis* (the basis in which $\hat{H}$ is diagonal). 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 $\hat{U}_H^\dagger$.

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{U}_H^\dagger(\boldsymbol{r},t) = \frac{\partial}{\partial t}\hat{U}_H^\dagger(\boldsymbol{r},t) + \boldsymbol{v}\cdot\nabla\hat{U}_H^\dagger(\boldsymbol{r},t), \tag{7.40}$$

where $\boldsymbol{v} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}$. Thus (7.38) becomes:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\chi_H(t)\rangle = \left[\hat{V}(\boldsymbol{r},t) - i\hbar\hat{U}_H(\boldsymbol{r},t)\frac{\partial}{\partial t}\hat{U}_H^\dagger(\boldsymbol{r},t) - i\hbar\boldsymbol{v}\cdot\hat{U}_H(\boldsymbol{r},t)\nabla\hat{U}_H^\dagger(\boldsymbol{r},t)\right]|\chi_H(t)\rangle. \tag{7.41}$$

The final term is identical to the non-adiabatic coupling term in the equation of motion as usually written in the surface-hopping literature [7] being a matrix with elements (in the eigenbasis):

$$\left(-i\hbar\boldsymbol{v}\cdot U_H(\boldsymbol{r},t)\nabla U_H^\dagger(\boldsymbol{r},t)\right)_{ij} = -i\hbar\boldsymbol{v}\cdot\langle\chi_i(\boldsymbol{r},t)|\nabla|\chi_j(\boldsymbol{r},t)\rangle, \tag{7.42}$$

[9] This is very similar to an interaction picture state vector (Section ??), 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.

where $U_H(\boldsymbol{r}, t)$ is the matrix representation of $\hat{U}$ in any basis that does not vary spatially (i.e, not the eigenbasis), $|\chi_i(\boldsymbol{r}, t)\rangle$ is the ith eigenvector of $\hat{H}(\boldsymbol{r}, t)$, and $\langle \chi_i(\boldsymbol{r}, t)|\nabla|\chi_j(\boldsymbol{r}, t)\rangle$ is the *non-adiabatic coupling vector* between the ith and j^{th} states referred to in the literature. The second to last term in brackets in (7.41) 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:

$$\frac{\mathrm{d}}{\mathrm{d}t}|c_i(t)|^2 = c_i(t)\frac{\mathrm{d}}{\mathrm{d}t}c_i^*(t) + c_i^*(t)\frac{\mathrm{d}}{\mathrm{d}t}c_i(t), \tag{7.43}$$

where via (7.41) we have:

$$\frac{\mathrm{d}}{\mathrm{d}t}c_i(t) = -\frac{\mathrm{i}}{\hbar}\sum_j (H_{\text{eff}}(\boldsymbol{r}, t))_{ij}\,c_j(t) \tag{7.44}$$

$$\Rightarrow \frac{\mathrm{d}}{\mathrm{d}t}c_i(t) = -\frac{\mathrm{i}}{\hbar}\sum_j \left[V_{ij}(\boldsymbol{r}, t) - \mathrm{i}\hbar\,\langle\chi_i(\boldsymbol{r}, t)|(\partial_t + \boldsymbol{v}\cdot\nabla)|\chi_j(\boldsymbol{r}, t)\rangle\right]c_j(t). \tag{7.45}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}t}|c_i(t)|^2 = \left[-\frac{\mathrm{i}}{\hbar}\sum_j c_i^*(t)\,(H_{\text{eff}}(\boldsymbol{r}, t))_{ij}\,c_j(t)\right] + \text{c.c.} \tag{7.46}$$

$$= -\frac{2}{\hbar}\sum_j \mathrm{Im}\left(c_i^*(t)\,(H_{\text{eff}}(\boldsymbol{r}, t))_{ij}\,c_j(t)\right) \tag{7.47}$$

$$= -\frac{2}{\hbar}\sum_j \mathrm{Im}\left(c_i^*(t)\left[V_{ij}(\boldsymbol{r}, t) - \mathrm{i}\hbar\,\langle\chi_i(\boldsymbol{r}, t)|(\partial_t + \boldsymbol{v}\cdot\nabla)|\chi_j(\boldsymbol{r}, t)\rangle\right]c_j(t)\right). \tag{7.48}$$

[10]This 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 $V(\boldsymbol{r}, t)$ is diagonal[10] and real, $c_i^*(t)V_{ij}(\boldsymbol{r}, t)c_j(t)$ is zero when $i \neq j$, and has no imaginary part when $i = j$, leaving us with

$$\frac{\mathrm{d}}{\mathrm{d}t}|c_i(t)|^2 = 2\sum_j \mathrm{Re}\left(c_i^*(t)\,\langle\chi_i(\boldsymbol{r}, t)|(\partial_t + \boldsymbol{v}\cdot\nabla)|\chi_j(\boldsymbol{r}, t)\rangle\,c_j(t)\right). \tag{7.49}$$

The change in $|c_i(t)|^2$ in a small interval $\mathrm{d}t$ is then:

$$|c_i(t + \mathrm{d}t)|^2 - |c_i(t)|^2 = 2\mathrm{d}t\sum_j \mathrm{Re}\left(c_i^*(t)\,\langle\chi_i(\boldsymbol{r}, t)|(\partial_t + \boldsymbol{v}\cdot\nabla)|\chi_j(\boldsymbol{r}, t)\rangle\,c_j(t)\right). \tag{7.50}$$

This is the change in the probability of the atom being in the ith 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 jth state to the ith 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

$$P_{ij}^{\text{space}} = \min\left\{q_{ij}^{\text{total}}, q_{ij}^{\text{space}}\right\}, \tag{7.51}$$

$$P_{ij}^{\text{time}} = \min\left\{q_{ij}^{\text{total}}, q_{ij}^{\text{time}}\right\}, \tag{7.52}$$

where

$$q_{ij}^{\text{space}} = 2\mathrm{d}t \, \mathrm{Re}\left(c_i^*(t) \, \langle \chi_i(\boldsymbol{r}, t)|\boldsymbol{v} \cdot \nabla|\chi_j(\boldsymbol{r}, t)\rangle \, c_j(t) \right), \tag{7.53}$$

$$q_{ij}^{\text{time}} = 2\mathrm{d}t \, \mathrm{Re}\left(c_i^*(t) \, \langle \chi_i(\boldsymbol{r}, t)|\partial_t|\chi_j(\boldsymbol{r}, t)\rangle \,, c_j(t) \right), \tag{7.54}$$

$$q_{ij}^{\text{total}} = \max\left\{ 0, (q_{ij}^{\text{space}} + q_{ij}^{\text{time}}) \right\}, \tag{7.55}$$

for transitions of the hidden variable from the jth to the ith eigenstate of $\hat{H}(\boldsymbol{r}, t)$ during the time interval $\mathrm{d}t$ due to non-adiabatic spatial and temporal variations in $\hat{H}(\boldsymbol{r}, t)$ respectively. These expressions can also be numerically integrated with respect to time to obtain transition probabilities over finite time intervals—numerically integrating over a short time interval Δt using a midpoint or higher order method can produce transition probabilities more accurate than $\mathcal{O}(\Delta t)$, which would be the accuracy if Δt were simply used in place of $\mathrm{d}t$ in the above expressions.

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:

$$P_{ii}^{\text{stay}} = |c_i(t)|^2 - \sum_{j \neq i} \left(P_{ij}^{\text{space}} + P_{ij}^{\text{time}} \right). \tag{7.56}$$

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 $\hat{H}$:

$$P = P^{\text{space}} + P^{\text{time}} + P^{\text{stay}}, \tag{7.57}$$

$$\Rightarrow \sum_j P_{ij} = |c_i(t)|^2, \tag{7.58}$$

$$\sum_i P_{ij} = |c_j(t + \mathrm{d}t)|^2. \tag{7.59}$$

The corresponding conditional probabilities of a transition to the ith state occurring—given that the hidden variable was already in the jth state—can be obtained via (7.16) as:

$$S_{ij}^{\text{space}} = \frac{1}{|c_j(t)|^2} P_{ij}^{\text{space}}, \tag{7.60}$$

$$S_{ij}^{\text{time}} = \frac{1}{|c_j(t)|^2} P_{ij}^{\text{time}} \tag{7.61}$$

$$S_{ij}^{\text{stay}} = \frac{1}{|c_j(t)|^2} P_{ij}^{\text{stay}}, \tag{7.62}$$

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

$$S = S^{\text{space}} + S^{\text{time}} + S^{\text{stay}}. \tag{7.63}$$

In practice we don't need to use this stochastic matrix to make transitions, 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.

[11] One can see that the $i = j$ term in (7.50) 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.

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 7.4).

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? As written, 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 (7.39) via

$$\hat{U}(t + dt, t) = e^{-\frac{i}{\hbar}\hat{H}_{\text{eff}}(\boldsymbol{r}, t)\, dt}, \tag{7.64}$$

where

$$\hat{H}_{\text{eff}} = \hat{V}(\boldsymbol{r}, t) - i\hbar \hat{U}_H(\boldsymbol{r}, t)\frac{d}{dt}U_H^\dagger(\boldsymbol{r}, t) \tag{7.65}$$

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:

$$H_{\text{eff}}(\boldsymbol{r}, t)\, dt = i\hbar \, \text{Log}\, U(t + dt, t) \tag{7.66}$$

$$\approx i\hbar\, (\mathbb{I} - U(t + dt, t)) \tag{7.67}$$

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

Since only the matrix $H_{\text{eff}}(\boldsymbol{r}, t)\, dt$ is required to compute transition probabilities according to (7.47), 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:

[12] Though we do need its component terms if we wish to distinguish between spatially vs. temporally induced transitions as in Section 7.4.2.

$$P_{ij} = \begin{cases} \max\left\{0, 2\,\text{Re}\left(c_i^*(t)U_{ij}(t + dt, t)c_j(t)\right)\right\} & i \neq j \\ |c_i(t)|^2 - \sum_{j \neq i} P_{ij} & i = j \end{cases}. \tag{7.68}$$

The corresponding stochastic matrix S can then be obtained by scaling the columns of P by the state populations as in (7.16). 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
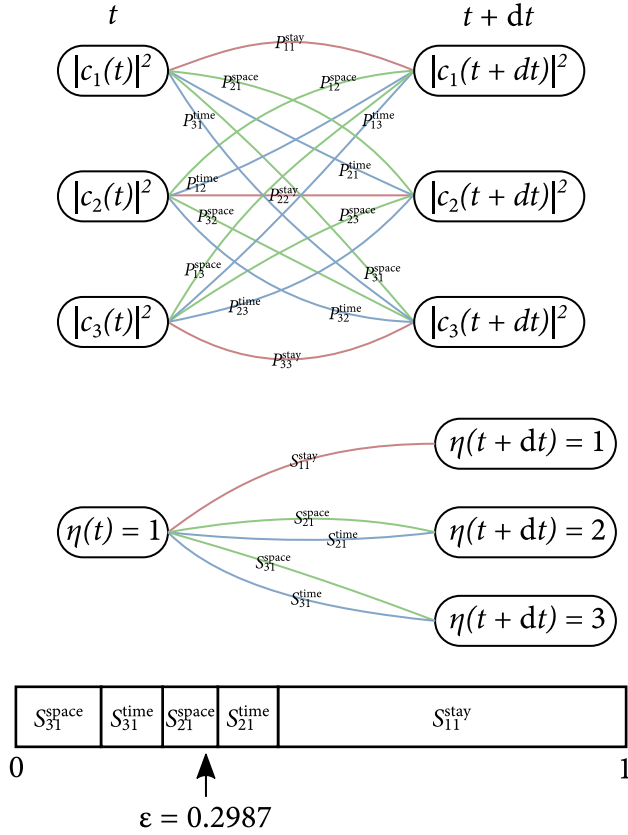
Figure 7.4: Probabilistically choosing a transition. Top: Probability flows between states in a time interval according to the three matrices of unconditional probabilities P^{stay}, P^{space}, and P^{time} in such a way that one total unit of probability is routed from states at the initial time to 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 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 spatially induced non-adiabatic transition.

for the given time interval is quite 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 $c_j(t)$, along with a factor of 2), then takes the real part of the result and clips it to zero from below. The two methods are not identical, but the similarity is striking.

7.4.3 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 jth state to the ith state due to the spatial non-adiabatic coupling term in the effective Hamiltonian is in the direction of the non-adiabatic coupling vector [7]

$$\boldsymbol{d}_{ij} = \langle \chi_i(\boldsymbol{r}, t) | \nabla | \chi_j(\boldsymbol{r}, t) \rangle \tag{7.69}$$

$$= \left(U_H(\boldsymbol{r}, t) \nabla U_H^\dagger(\boldsymbol{r}, t) \right)_{ij} \tag{7.70}$$

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

$$\Delta v_{ij}^2 = \frac{2}{m} \left(V_j(\boldsymbol{r}, t) - V_i(\boldsymbol{r}, t) \right), \tag{7.71}$$

where $V_i(\boldsymbol{r}, t)$ is the adiabatic potential comprising the ith spatially and/or temporally varying eigenvalue of $\hat{H}(\boldsymbol{r}, t)$. 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 $\tilde{P}^{\text{space}}$ that sets the probability of the disallowed transitions to zero:

$$\tilde{P}_{ij}^{\text{space}} = \begin{cases} P_{ij}^{\text{space}} & \Delta v_{ij}^2 + |\boldsymbol{v} \cdot \hat{\boldsymbol{d}}_{ij}|^2 \geq 0 \\ 0 & \Delta v_{ij}^2 + |\boldsymbol{v} \cdot \hat{\boldsymbol{d}}_{ij}|^2 < 0 \end{cases} \tag{7.72}$$

where $\hat{\boldsymbol{d}}_{ij}$ is the unit vector in the direction of $\boldsymbol{d}_{ij}$.[13]

The diagonal matrix for the probabilities of remaining in each state must be adjusted similarly to absorb the probability discarded this way:

$$\tilde{P}_{ii}^{\text{stay}} = |c_i(t)|^2 - \sum_{j \neq i} \left(\tilde{P}_{ij}^{\text{space}} + P_{ij}^{\text{time}} \right) \tag{7.73}$$

The corresponding matrices of conditional probabilities for the hidden variable transitioning, given that it is already in the jth state, are

$$\tilde{S}_{ij}^{\text{space}} = \frac{1}{|c_j(t)|^2} \tilde{P}_{ij}^{\text{space}}, \tag{7.74}$$

$$\tilde{S}_{ij}^{\text{stay}} = \frac{1}{|c_j(t)|^2} \tilde{P}_{ij}^{\text{stay}}, \tag{7.75}$$

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 $\tilde{P}^{\text{space}}$, $\tilde{P}^{\text{time}}$, $\tilde{S}^{\text{space}}$, and $\tilde{S}^{\text{time}}$ do not need to be computed at each timestep.[14]

In the case that a transition of the hidden variable does occur to the ith from the jth 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:

$$\Delta \boldsymbol{v}_{ij} = \left[\text{sgn}(\boldsymbol{v} \cdot \hat{\boldsymbol{d}}_{ij}) \sqrt{(\boldsymbol{v} \cdot \hat{\boldsymbol{d}}_{ij})^2 + \frac{2}{m} \left(V_i(\boldsymbol{r}, t) - V_j(\boldsymbol{r}, t) \right)} - (\boldsymbol{v} \cdot \hat{\boldsymbol{d}}_{ij}). \right] \hat{\boldsymbol{d}}_{ij} \tag{7.76}$$

7.5 Decoherence

In the context of the hidden-variable semiclassical/surface hopping method, *decoherence* refers to the fact that, due to positional separation of different internal states of the atom,

[13] The surface hopping literature calls these classically disallowed transitions *frustrated* transitions. Prior to encountering the surface-hopping literature, my method of conserving energy was identical to this (for the case of time-independent potentials), with the exception that I did not know what direction the velocity kick ought to be in, limiting my method to one spatial dimension only. There is much interest in alternate methods of energy conservation in chemical physics, since bound electrons can spend quite a bit of their time energetically 'out of their depths' with negative kinetic energy. However this is outside of the scope of my interest in these methods for simulating cold atoms, where purely classical modelling of motional states is likely to be more acceptable than for bound electrons.

[14] Though one of my decoherence methods (Section 7.5.4) does make use of the full matrices in order to track transitions between other states for the purpose of computing approximate decoherence factors.

those internal states transition from being a coherent superposition into a statistical mixture. For example, in the Stern–Gerlach experiment, a spin may be initially a coherent superposition of spin-up and spin-down, but by the time the two components separate, the superposition is no longer a coherent one. The spin is either up (and the atom off to one one side of the screen) or down (and the atom over the other side of the screen). At this point, no interference between the two internal states is observable, as they are not co-located in space. The motional degree of freedom has played the role of an environment, and, having become entangled with the spin of the atom, decohered the spin state. Recoherence can occur if the two wavepackets are brought back together again to have well overlapping positions and velocities, but this is difficult to achieve even intentionally[15], let alone by accident,

Decoherence was not part of Tully's original surface hopping model [5], such that if used to simulate the Stern–Gerlach experiment, two spots would result on the screen, but both spots would comprise a coherent $50:50$ superposition of spin-up and spin-down, rather than one being up and one being down. Some decoherence is necessary to include in a hidden-variable semiclassical/surface hopping model in order to avoid this unphysical outcome, and furthermore, to obtain correct transition probabilities in the case that the atom undergoes multiple periods of local spin transitions (as being in a $50:50$ superposition represents entirely different initial conditions for a Majorana spin flip than being in a single spin state).

In Section 7.5.1 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. I then show in Section 7.5.2 why naively projecting the wavefunction onto a single trajectory at all times yields nonsensical results, and then in Subsections 7.5.3 and 7.5.3 I present two remedies for how to approximately include decoherence despite this.

[15] The problem of getting the wavepackets back together again has been coined the "Humpty-Dumpty problem" [5, 30] and has made magnetic separation impractical for large-momentum atom interferometry [31], and methods such as the use of Bragg pulses [32, 33] are needed 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.

7.5.1 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 $|r\rangle$:

$$|\Psi(t)\rangle = \int \sum_i \psi_i(r, t) |\chi_i\rangle \otimes |r\rangle \, dr, \qquad (7.77)$$

where normalisation requires that

$$\int \sum_i |\psi_i(r, t)|^2 dr = 1. \qquad (7.78)$$

Recognising that $\psi_i(r, t)$ is the ith 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 $\phi_i(r, t)$ and corresponding motional state vector $|\phi_i(t)\rangle$ and its coefficient $c_i(t)$ for each component:

$$c_i(t)\phi_i(r, t) \equiv \psi_i(r, t) \qquad (7.79)$$

$$|\phi_i(t)\rangle \equiv \int \phi_i(r, t) |r\rangle \, dr \qquad (7.80)$$

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

$$|\Psi(t)\rangle = \sum_i c_i(t) |\chi_i\rangle \otimes |\phi_i(t)\rangle, \qquad (7.81)$$

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 $|\Psi(t)\rangle$:

$$\hat{\rho}(t) = |\Psi(t)\rangle\langle\Psi(t)| = \sum_{ij} c_i(t)c_j^*(t) |\chi_i(t)\rangle\langle\chi_j(t)| \otimes |\phi_i(t)\rangle\langle\phi_j(t)| , \tag{7.82}$$

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

$$\rho_{ij}(t, \boldsymbol{r}, \boldsymbol{r}') = \langle \chi_i\, \boldsymbol{r} | \rho(t) | \chi_j\, \boldsymbol{r}' \rangle , \tag{7.83}$$

$$= \psi_i(\boldsymbol{r}, t)\psi_j^*(\boldsymbol{r}', t), \tag{7.84}$$

$$= c_i(t)c_j^*(t)\phi_i(\boldsymbol{r}, t)\phi_j^*(\boldsymbol{r}', t). \tag{7.85}$$

A partial trace [34] 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:

$$\hat{\rho}^{\text{red}}(t) = \int \langle \boldsymbol{r} | \hat{\rho}(t)\hat{r} | \boldsymbol{r} \rangle \, \mathrm{d}\boldsymbol{r} \tag{7.86}$$

$$\Rightarrow \rho_{ij}^{\text{red}}(t) = \int \rho_{ij}(t, \boldsymbol{r}, \boldsymbol{r}')\delta(\boldsymbol{r} - \boldsymbol{r}')\mathrm{d}\boldsymbol{r} \tag{7.87}$$

$$= c_i(t)c_j^*(t) \int \phi_i(\boldsymbol{r}, t)\phi_j^*(\boldsymbol{r}, t)\mathrm{d}\boldsymbol{r} \tag{7.88}$$

$$= c_i(t)c_j^*(t) \langle \phi_j(t)|\phi_i(t)\rangle . \tag{7.89}$$

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 7.3.

To obtain an explicit form for the decoherence factor $R_{ij}(t)$, we need to start imposing our assumptions about the motional states $\{ |\phi_i(t)\rangle \}$. As alluded to earlier, we will be assuming that each motional state $|\phi_i(t)\rangle$ is a Gaussian wavepacket propagating—without dispersion—with centre-of-mass motion evolving classically according to the adiabatic potential $V_i(\boldsymbol{r}, t)$ experienced by the local eigenstate $|\chi_i\rangle$. Thus the wavefunctions of these motional states are

$$\langle \boldsymbol{r} | \phi_i(t)\rangle = \phi_i(\boldsymbol{r}, t) = A \exp\left[-\frac{|\boldsymbol{r} - \boldsymbol{r}_i(t)|^2}{4\sigma^2} + \mathrm{i}\frac{m}{\hbar}\boldsymbol{v}_i(t) \cdot (\boldsymbol{r} - \boldsymbol{r}_i(t)) \right], \tag{7.90}$$

Where A is a real normalisation constant,[16] $\boldsymbol{r}_{ij}(t) = \boldsymbol{r}_i(t) - \boldsymbol{r}_j(t)$ and $\boldsymbol{k}_{ij}(t) = \frac{m}{\hbar}\left(\boldsymbol{v}_i(t) - \boldsymbol{v}_j(t) \right)$ are the relative mean position and wavevector of the pair of wavepackets, and where the centre-of-mass position and velocity of each wavepacket evolve classically according to

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{r}_i(t) = \boldsymbol{v}_i(t) \tag{7.91}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{v}_i(t) = -\frac{1}{m}\nabla V_i(\boldsymbol{r}_i, t). \tag{7.92}$$

[16] The overall phase is determined by the offset $\boldsymbol{r} - \boldsymbol{r}_i(t)$ in the second term in the exponent, and is chosen such that $\phi_i(\boldsymbol{r}, t)$ is real at $\boldsymbol{r} = \boldsymbol{r}_i$, 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 later apply our reduction operator $\hat{R}$ would not affect the system dynamics, but is unappealing.

This gives for the decoherence factor:

$$R_{ij}(t) = \exp\left[-\left(\frac{1}{8\sigma^2}|\boldsymbol{r}_{ij}|^2 + \frac{i}{2}\boldsymbol{r}_{ij}(t)\cdot\boldsymbol{k}_{ij}(t) + \frac{\sigma^2}{2}|\boldsymbol{k}_{ij}(t)|^2\right)\right]. \tag{7.93}$$

As expected, $R_{ij}(t)$ is equal to unity when the two wavepackets have identical positions and velocities, and decays to zero for increasing relative position and velocity.

If the two motional states being considered are identical at $t = 0$ and have constant relative acceleration $\boldsymbol{a}_{ij}$, then we have

$$\boldsymbol{r}_{ij}(t) = \frac{1}{2}\boldsymbol{a}_{ij}t^2 \tag{7.94}$$

and

$$\boldsymbol{k}_{ij}(t) = \frac{m}{\hbar}\boldsymbol{a}_{ij}t, \tag{7.95}$$

which gives for the decoherence factor:

$$R_{ij}(t) = \exp\left[-\frac{|\boldsymbol{a}_{ij}|^2}{2}\left(\frac{1}{16\sigma^2}t^4 + i\frac{m}{2\hbar}t^3 + \frac{m^2\sigma^2}{\hbar^2}t^2\right)\right]. \tag{7.96}$$

We will return later to this explicit form of the decoherence factor under these assumptions.

Returning to our arbitrary state vector (7.79), we find (ignoring normalisation) the projected state vector $|\tilde{\Psi}(t)\rangle = \hat{R}(t)|\tilde{\Psi}(t)\rangle$, 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.

$$|\tilde{\Psi}(t)\rangle = \hat{R}(t)|\Psi(t)\rangle \tag{7.97}$$

$$= \sum_i \langle\phi_\eta(t)|\phi_i(t)\rangle c_i(t)|\chi_i\rangle \otimes |\phi_\eta(t)\rangle \tag{7.98}$$

$$= \sum_i \tilde{c}_i(t)|\chi_i\rangle \otimes |\phi_\eta(t)\rangle \tag{7.99}$$

$$= |\tilde{\chi}(t)\rangle \otimes |\phi_\eta(t)\rangle, \tag{7.100}$$

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:

$$|\tilde{\Psi}(t+dt)\rangle = \hat{R}(t+dt)\sum_i\left[\tilde{c}_i(t) - \frac{i}{\hbar}\sum_j H_{ij}(t)\tilde{c}_j(t)dt\right]|\chi_i\rangle \otimes |\phi_i(t+dt)\rangle, \tag{7.101}$$

where $H_{ij}(t) = \langle\chi_i|\hat{H}_\eta(t)|\chi_j\rangle = \langle\chi_i\,\phi_\eta(t)|\hat{H}(t)|\chi_j\,\phi_\eta(t)\rangle$ are the elements of the projected Hamiltonian $\hat{H}_\eta(t)$ 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 ith 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

$$\langle \phi_\eta(t+dt)|\phi_i(t+dt)\rangle = \langle \phi_\eta(t)|\phi_i(t)\rangle + \frac{d}{dt}\langle \phi_\eta(t)|\phi_i(t)\rangle\, dt \qquad (7.102)$$

$$= 1 + \frac{dR_{\eta i}(t)}{dt}. \qquad (7.103)$$

Using this fact in applying the projection in (7.101) yields a product state once more:

$$|\tilde{\Psi}(t+dt)\rangle = \sum_i \left(1 + \frac{dr_{\eta i}(t)}{dt}\right)\left[\tilde{c}_i(t) - \frac{i}{\hbar}\sum_j H_{ij}(t)\tilde{c}_j(t)dt\right]|\chi_i\rangle \otimes |\phi_\eta(t+dt)\rangle \qquad (7.104)$$

$$\Rightarrow |\tilde{\chi}(t+dt)\rangle = \sum_i \left(1 + \frac{dr_{\eta i}(t)}{dt}\right)\left[\tilde{c}_i(t) - \frac{i}{\hbar}\sum_j H_{ij}(t)\tilde{c}_j(t)dt\right]|\chi_i\rangle \qquad (7.105)$$

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

$$\frac{d}{dt}|\tilde{\chi}(t)\rangle = \left[-\frac{i}{\hbar}\hat{H}_\eta(t) - \hat{\Gamma}_\eta(t)\right]|\tilde{\chi}(t)\rangle, \qquad (7.106)$$

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

$$(\Gamma_\eta(t))_{ii} = \gamma_{\eta i}(t) = -\frac{dR_{\eta i}(t)}{dt} = -\frac{d}{dt}\langle \phi_\eta(t)|\phi_i(t)\rangle, \qquad (7.107)$$

where we have defined $\gamma_{ij}(t) = -\frac{dR_{ij}(t)}{dt}$. The diagonals of $\hat{\Gamma}_\eta(t)$ in the eigenbasis 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 $|\phi_\eta(t)\rangle$.

7.5.2 The quantum Zeno effect

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 zero at all times, and so our decoherence rates are:

$$\gamma_{ij} = -\left[\frac{dR_{ij}(t)}{dt}\right]_{\substack{k_{ij}=0 \\ \nu_{ij}=0}} \qquad (7.108)$$

$$= -\left[\frac{\partial R_{ij}(t)}{\partial r_{ij}}\cdot\frac{dr_{ij}(t)}{dt} + \frac{\partial R_{ij}(t)}{\partial k_{ij}}\cdot\frac{dk_{ij}(t)}{dt}\right]_{\substack{k_{ij}=0 \\ \nu_{ij}=0}} \qquad (7.109)$$

$$= \left[\left(\frac{r_{ij}}{4\sigma^2} + \frac{i}{2}k_{ij}\right)R_{ij}(t)\cdot\frac{dr_{ij}(t)}{dt} + \left(\sigma^2 k_{ij} + \frac{i}{2}r_{ij}\right)R_{ij}(t)\cdot\frac{dk_{ij}(t)}{dt}\right]_{\substack{k_{ij}=0 \\ \nu_{ij}=0}} \qquad (7.110)$$

$$= 0. \qquad (7.111)$$

Every decoherence rate is zero. What is going on? The answer is the quantum Zeno effect [35, 36], which is the name given to the fact that, in the 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. In our case, the fact that we are projecting constantly onto our assumed motional state causes the amplitude flows to the other motional states to be exactly zero. A sufficiently closely watched quantum pot never boils, and a sufficiently closely watched Schrödinger's cat never may never be poisoned [36]:

> In view of the Zeno's paradox formulated above, should we conclude that the particle will never decay? Will the cat escape the cruel death awaiting it, against which it has no defense, provided its vital signs are constantly watched with loving care?

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, if possible with a differential equation that does not require us to simulate all the quantum details of the motional degree of freedom.

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

$$R_{ij}^{\text{Markov}}(t) = e^{-\gamma_{ij}t}, \tag{7.112}$$

then the decoherence rate would be the constant γ_{ij}, and not zero. This is the case for Markovian decoherence [34], which is when the environment has no memory of its past interaction with the system. The memory in our case is due to the wavepackets accelerating 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 [37]. 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 [18, 19, 38]. 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 [19]. The measurement interval assumed "should be large enough to allow the photons to get away from the atom" [39]. This can be recast as a continuous *weak* measurement, rather than infrequent projective measurements [19, 40], 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 $\frac{\mathrm{d}\tilde{c}_i(t)}{\mathrm{d}t} = -\gamma \tilde{c}_i(t)$ 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. The first, described in Section 7.5.3, is to gloss over as many of the details of the wavepacket separation as possible and replace the decoherence factor with an exponential one, describing the wavepackets separating on approximately the correct timescale. This is crude, but better than no decoherence at all. The second, described in Section 7.5.4, is to introduce a minimal memory of the separation of wavepackets. In this approach, one (but only one) trajectory is simulated for the motional state corresponding to each internal state. Whenever there is population transfer between states, these trajectories are replaced with a weighted average of their existing position with that of the source of the population transfer.

7.5.3   Approximate Markovian decoherence

A crude way to include decoherence is just to approximate $R_{\eta i}(t)$ as a decaying exponential with roughly the right decay constant. This is crude because $R_{\eta i}(t)$ does not look much like a decaying exponential (see Figure 7.5). In the limit of large t, for the case constant acceleration as in (7.96) its functional form is e^{-t^4}, not the exponential decay required to treat the decoherence as Markovian at any timescale.

To nonetheless find an approximate Markovian decoherence rate, we first construct a "time ignorant" version $\tilde{R}_{ij}(t)$ of the decoherence factor $\tilde{R}_{ij}(t)$ given in (7.96) that answers the question "What is the expected value of $R_{ij}(t)$ 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 $\gamma_{ij}^{\text{Markov}}$. 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 [5], which also lacked any decoherence. Others [41–43] have since developed various methods to induce approximate damping of states to achieve a similar outcome, including the approximation of an exponential decoherence factor [41], though my development of this method was independent, as I was unaware of this literature at the time.

Proceeding, we define the time-ignorant decoherence factor $\tilde{R}_{ij}(t)$ as the average[17] of all decoherence factors one would obtain if the two wavepackets began separating at some point in time before $t = 0$:

[17] Not a quantum expectation value, just a weighted sum.

$$\tilde{R}_{ij}(t) = A \int_{-\infty}^{0} \langle \phi_i(t - t') | \phi_j(t - t') \rangle \; \mathrm{d}t' \tag{7.113}$$

$$= A \int_{-\infty}^{0} R_{ij}(t - t') \, \mathrm{d}t' \tag{7.114}$$

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 with constant relative acceleration as in (7.91) and (7.92) such that the $R_{ij}(t)$ above has the form (7.96).

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

$$\gamma_{ij}^{\text{Markov}} = -\frac{\left[\frac{\mathrm{d}}{\mathrm{d}t} \int_{-\infty}^{0} R_{ij}(t - t') \, \mathrm{d}t' \right]_{t=0}}{\left[\int_{-\infty}^{0} R_{ij}(t - t') \, \mathrm{d}t' \right]_{t=0}}. \tag{7.115}$$
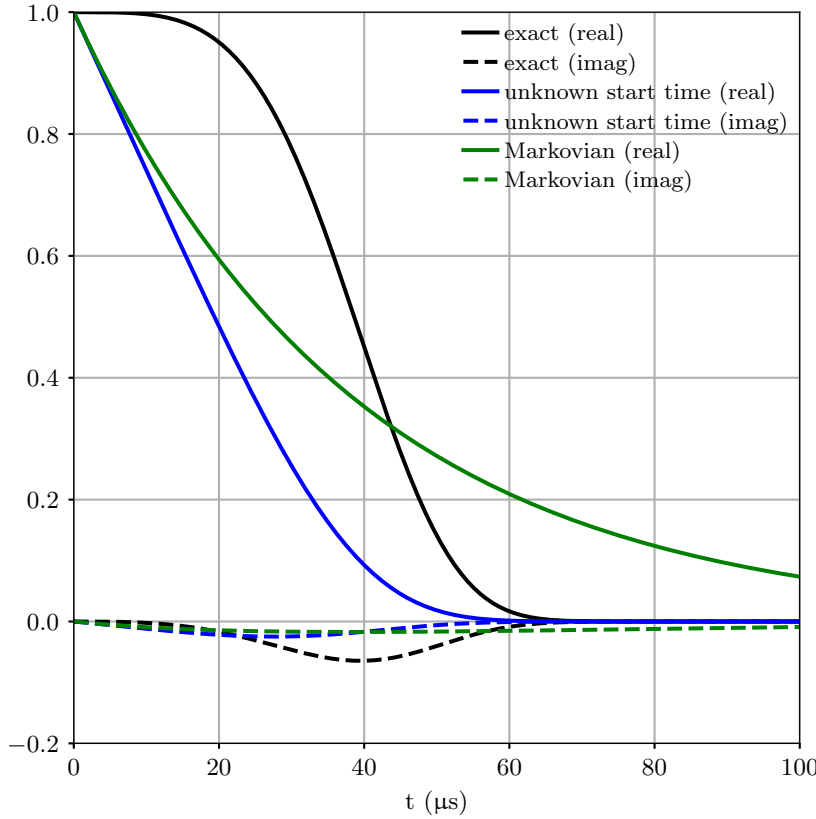
Figure 7.5: Example decoherence factor as a function of time for an experimentally realistic parameters. Plotted in black is the exact decoherence factor $R_{ij}(t)$ (equation (7.96)) between two adjacent ($\Delta m_F = \pm 1$) Zeeman sublevels of the $F = 1$ groundstate of ^{87}Rb, assuming constant relative acceleration due to a magnetic field gradient of $250\,\mathrm{G\,cm^{-1}}$ and a Gaissian wavepacket width $\sigma = \lambda_{\mathrm{th}} = 2.6\,\mu\mathrm{m}$, corresponding to the thermal wavelength $\lambda_{\mathrm{th}} = h/\sqrt{2\pi m k_B T}$ at $T = 53\,\mu\mathrm{K}$. Shown in blue is the 'time-ignorant' decoherence factor described in text. In green is the Markovian approximation to the exact decoherence factor, obtained by extracting a decay constant from the gradient of the time-ignorant decoherence factor at $t = 0$. As can be seen, whilst no decaying exponential is a particularly good fit to the exact decoherence factor, our Markovian approximation is as good as can be expected, decaying to zero over approximately the correct timescale.

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

$$\gamma_{ij}^{\mathrm{Markov}} = -\frac{\int_{-\infty}^{0} R'_{ij}(-t')\,\mathrm{d}t'}{\int_{-\infty}^{0} R_{ij}(-t')\,\mathrm{d}t'} \tag{7.116}$$

$$= -\frac{\int_{0}^{\infty} R'_{ij}(t')\,\mathrm{d}t'}{\int_{0}^{\infty} R_{ij}(t')\,\mathrm{d}t'} \tag{7.117}$$

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

$$\frac{1}{\gamma_{ij}^{\text{Markov}}} = \int_0^\infty R_{ij}(t')\,dt' \tag{7.118}$$

$$= \int_0^\infty \exp\left[-\left(\frac{1}{8\sigma^2}|\boldsymbol{r}_{ij}|^2 + \frac{i}{2}\boldsymbol{r}_{ij}(t')\cdot\boldsymbol{k}_{ij}(t') + \frac{\sigma^2}{2}|\boldsymbol{k}_{ij}(t')|^2\right)\right]dt' \tag{7.119}$$

The expression (7.96) requires a relative acceleration, for which we use the relative acceleration between the pair of adiabatic potentials at the current moment in time and current position of the atom during a simulation:

$$\boldsymbol{a}_{ij}(\boldsymbol{r},t) \approx -\frac{1}{m}\left(\nabla V_i(\boldsymbol{r},t) - \nabla V_j(\boldsymbol{r},t)\right). \tag{7.120}$$

This now gives a decoherence factor depending on position and time:

$$\frac{1}{\gamma_{ij}^{\text{Markov}}(\boldsymbol{r},t)} = \int_0^\infty \exp\left[-\frac{|\boldsymbol{a}_{ij}(\boldsymbol{r},t)|^2}{2}\left(\frac{1}{16\sigma^2}t'^4 + i\frac{m}{2\hbar}t'^3 + \frac{m^2\sigma^2}{\hbar^2}t'^2\right)\right]dt' \tag{7.121}$$

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 (7.118) 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:

$$\frac{1}{\gamma_{ij\,(\text{pos})}^{\text{Markov}}(\boldsymbol{r},t)} \approx \int_0^\infty \exp\left[-\frac{|\boldsymbol{a}_{ij}(\boldsymbol{r},t)|^2}{2}\left(\frac{1}{16\sigma^2}t'^4 + i\frac{m}{2\hbar}t'^3\right)\right]dt' \tag{7.122}$$

$$\approx \int_0^\infty \exp\left[-\frac{|\boldsymbol{a}_{ij}(\boldsymbol{r},t)|^2}{32\sigma^2}t'^4\right]\left(1 - i\frac{m|\boldsymbol{a}_{ij}(\boldsymbol{r},t)|^2}{4\hbar}t'^3\right)dt' \tag{7.123}$$

$$= 2^{\frac{5}{4}}\Gamma(\tfrac{5}{4})\sqrt{\frac{\sigma}{|\boldsymbol{a}_{ij}(\boldsymbol{r},t)|}} - 2i\frac{m\sigma^2}{\hbar}, \tag{7.124}$$

where we used a first-order Taylor expansion of an exponential in (7.123). We similarly use a first-order expansion $(x+\varepsilon)^{-1} \approx x^{-1} - \varepsilon x^{-2}$ to take the reciprocal of (7.124) (since the second term is much smaller than the first[18]), and arrive at:

$$\gamma_{ij\,(\text{pos})}^{\text{Markov}}(\boldsymbol{r},t) \approx \frac{1}{2^{\frac{5}{4}}\Gamma(\tfrac{5}{4})}\sqrt{\frac{|\boldsymbol{a}_{ij}(\boldsymbol{r},t)|}{\sigma}} + \frac{i}{2^{\frac{3}{2}}\Gamma(\tfrac{5}{4})^2}\frac{m\sigma|\boldsymbol{a}_{ij}(\boldsymbol{r},t)|}{\hbar}. \tag{7.125}$$

[18] This 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.

Similarly for the large σ regime, we neglect the first term in the exponent of (7.118) and consider the second term small relative to the third. This is the regime in which the decrease in overlap of the two wavepackets is dominated by their separation in velocity

space. Following the same process as above gives:

$$\frac{1}{\gamma_{ij\,(\text{vel})}^{\text{Markov}}(\boldsymbol{r},t)} \approx \int_0^\infty \exp\left[-\frac{|\boldsymbol{a}_{ij}(\boldsymbol{r},t)|^2}{2}\left(\mathrm{i}\frac{m}{2\hbar}t'^3 + \frac{m^2\sigma^2}{\hbar^2}t'^2\right)\right]\mathrm{d}t' \qquad (7.126)$$

$$\approx \int_0^\infty \left(1 - \mathrm{i}\frac{m|\boldsymbol{a}_{ij}(\boldsymbol{r},t)|^2}{4\hbar}t'^3\right)\exp\left[-\frac{m^2\sigma^2|\boldsymbol{a}_{ij}(\boldsymbol{r},t)|^2}{2\hbar^2}t'^2\right]\mathrm{d}t$$
$$(7.127)$$

$$= \sqrt{\frac{\pi}{2}}\frac{\hbar}{m\sigma|\boldsymbol{a}_{ij}(\boldsymbol{r},t)|} - \mathrm{i}\frac{\hbar^3}{2m^3\sigma^4|\boldsymbol{a}_{ij}(\boldsymbol{r},t)|^2} \qquad (7.128)$$

$$\Rightarrow \gamma_{ij\,(\text{vel})}^{\text{Markov}}(\boldsymbol{r},t) \approx \sqrt{\frac{2}{\pi}}\frac{m\sigma|\boldsymbol{a}_{ij}(\boldsymbol{r},t)|}{\hbar} + \mathrm{i}\frac{\hbar}{\pi m\sigma^2} \qquad (7.129)$$

Equations (7.125) and (7.129) are our final expressions for the Markovian docoherence 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 $\gamma_{ij}^{\text{Markov}}(\boldsymbol{r},t)$ over all wavepacket sizes:

$$\gamma_{ij}^{\text{Markov}}(\boldsymbol{r},t) \approx \left[\text{Re}(\gamma_{ij\,(\text{pos})}^{\text{Markov}}(\boldsymbol{r},t))^2 + \text{Re}(\gamma_{ij\,(\text{vel})}^{\text{Markov}}(\boldsymbol{r},t))^2\right]^{\frac{1}{2}}$$
$$+ \mathrm{i}\left[\text{Im}(\gamma_{ij\,(\text{pos})}^{\text{Markov}}(\boldsymbol{r},t))^{-1} + \text{Im}(\gamma_{ij\,(\text{vel})}^{\text{Markov}}(\boldsymbol{r},t))^{-1}\right]^{-1}. \qquad (7.130)$$

We now have an approximate analytic expression for a Markovian decoherence rate 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 (7.130), compared to the exact expression (7.121) for $\gamma_{ij}^{\text{Markov}}$ over a range of wavepacket sizes is shown in Figure 7.6.

7.5.4 Decoherence with mean auxiliary trajectories

A more accurate approach to including decoherence is to compute a decoherence factor at each timestep of the simulation by modelling the wavepackets as they accelerate away from each other, retaining in the model the position and velocity of each wavepacket. This inclusion of additional positions and velocities in the model comprises a minimal memory of the past interaction between the system and environment—that is, the internal and motional degrees of freedom of the atom.

In quantum mechanics there is not, in general, a single Gaussian wavepacket for each state, and even if the wavefunction corresponding to the originating state is a Guassian wavepacket, transitions to other states at different times can produce an arbitrary superposition of Gaussian wavepackets in each other state. If we were to simulate this arbitrary superposition for all components of the multi-component wavefunction, we would not be saving any computational power at all, as we would have reverted to a fully quantum description of the motional degree of freedom, which is not at all our intention. Therefore we draw the line at modelling a single trajectory for each state, in order to maintain simplicity whilst including at least some dynamics of the system–environment interaction.

Below I will talk about how to convert a decoherence factor at each point in time into a decoherence *rate* at each point in time, that can be included in a differential equation for the state vector of the atom's internal state. Then I will introduce the averaging scheme I am using in order to turn what would be many trajectories for each state into one average trajectory.

This use of 'auxiliary trajectories' for computing decoherence is common in the surface-hopping literature. There is a range of methods called 'multiple spawning' methods, in
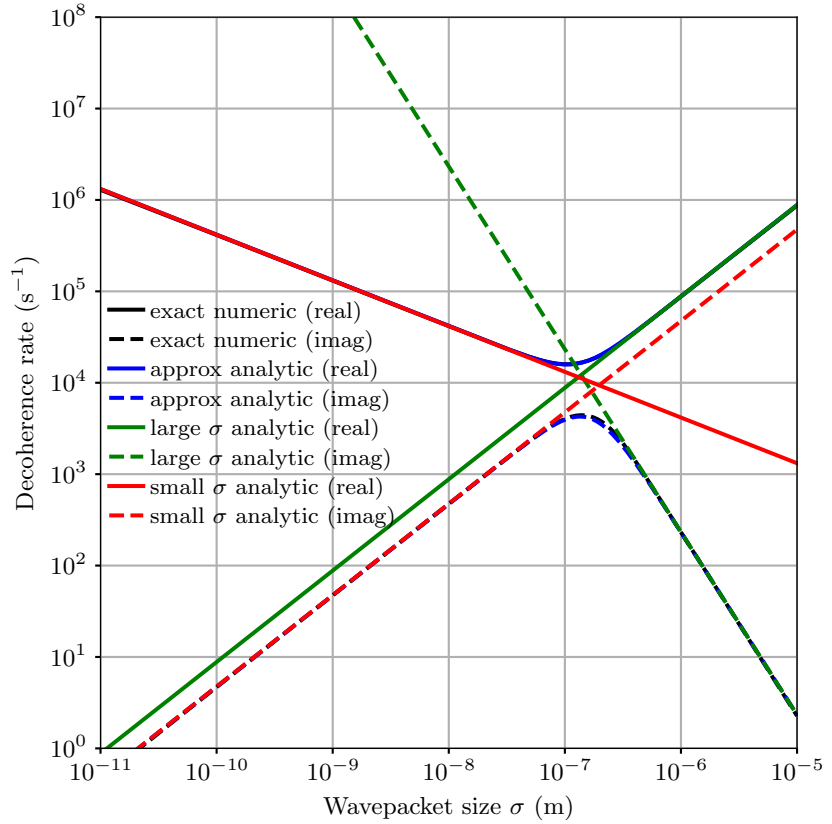
Figure 7.6: Comparison of approximate analytic Markovian decoherence rate with exact Markovian decoherence rate between adjacent ($\Delta m_F = \pm 1$) Zeeman sublevels of the $F = 1$ groundstate of a ^{87}Rb atom in a $250\,\text{G}\,\text{cm}^{-1}$ magnetic field gradient. Real (solid) and imaginary (dashed) parts of the small wavepacket (red) and large wavepacket (green) limits are shown, their stitching together in the intermediate wavepacket regime (blue) and the exact expression they approximate (black, not easily visible due to being similar to the blue curves). There is good agreement over all wavepacket sizes.

which additional auxiliary trajectories regularly branch off the main trajectory (the one corresponding to the hidden variable), and are all considered in order to compute the overlap with the main wavepacket and hence the decoherence factor [44–46]. These methods generally contain a parameter controlling the likelihood of new trajectories branching off, and trajectories can be discarded once they recede far enough from the main trajectory. If the parameter is tuned far enough, there can be so many auxiliary trajectories that an extremely accurate decoherence rate can be computed, though the computational cost in this case approaches that of a fully quantum simulation. A method by Shenvi et al. [9], like mine, tracks only one auxiliary trajectory per internal atomic state, probabilistically replacing them with newly spawned trajectories rather than simulating multiple auxiliary trajectories per state. However, the trajectories are spawned at specific moments—when the non-adiabatic coupling strength reaches a local maximum—leading to different behaviour in different locations in space. For example, if an atom were orbiting a magnetic field zero in a magnetic trap, this method would only spawn new trajectories at the orbit's point of closest approach, and decoherence rates would be inaccurate at other points in the orbit, even if the non-adiabatic coupling were still high throughout

the orbit. My method, though not thoroughly tested, has the potential to improve upon this due to the fact that it is in a sense spawning new trajectories all the time—rather than only at specific points in space as the method by Shenvi et al.—and avoids the problem of an exponential proliferation of trajectories by only retaining an average of each state's trajectory rather than a set of them.

Decoherence rate in the absence of projective measurement

Here we generalise our conception of the decoherence rate by removing the assumption that motional states are "reset" at each timestep. This resetting, which we assumed in Section 7.5.1, set each motional state to be equal to $|\phi_\eta(t)\rangle$ at each timestep, and this led to all decoherence rates being zero due to the quantum Zeno effect (Section 7.5.2). What is the reduction in the state amplitudes at each timestep if the wavepackets are 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 ith state amplitude would be reduced by a factor of $R_{\eta i}(t)$ compared to the original state amplitude $c_i(t)$:

$$\tilde{c}_i(t) = R_{\eta i}(t) c_i(t). \tag{7.131}$$

This implies:

$$\frac{\mathrm{d}\tilde{c}_i(t)}{\mathrm{d}t} = \frac{\mathrm{d}R_{\eta i}(t)}{\mathrm{d}t} c_i(t) \tag{7.132}$$

$$= \frac{1}{R_{\eta i}(t)} \frac{\mathrm{d}R_{\eta i}(t)}{\mathrm{d}t} \tilde{c}_i(t). \tag{7.133}$$

and so we see that a a decoherence rate that does not reset the environment states is the (negative of the) logarithmic derivative of the decoherence factor at any time, rather than just its derivative at $t = 0$. 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. Thus, given the decoherence factor between two states, along with its time derivative, we can compute a decoherence rate

$$\gamma_{ij}^{\mathrm{mm}}(t) = -\frac{1}{R_{ij}(t)} \frac{\mathrm{d}R_{ij}(t)}{\mathrm{d}t}, \tag{7.134}$$

such that in the absence of Hamiltonian evolution

$$\frac{\mathrm{d}\tilde{c}_i(t)}{\mathrm{d}t} = -\gamma_{\eta i}^{\mathrm{mm}}(t) \, \tilde{c}_i(t), \tag{7.135}$$

where "mm" stands for "minimal-memory".

The (negative) logarithmic derivative of the decoherence factor for our fixed-width Gaussian wavepackets (7.93) gives the decoherence rate

$$\gamma_{ij}^{\mathrm{mm}}(t) = \left(\frac{\boldsymbol{r}_{ij}}{4\sigma^2} + \frac{\mathrm{i}}{2} \boldsymbol{k}_{ij} \right) \cdot \frac{\mathrm{d}\boldsymbol{r}_{ij}(t)}{\mathrm{d}t} + \left(\sigma^2 \boldsymbol{k}_{ij} + \frac{\mathrm{i}}{2} \boldsymbol{r}_{ij} \right) \cdot \frac{\mathrm{d}\boldsymbol{k}_{ij}(t)}{\mathrm{d}t}, \tag{7.136}$$

which for a given relative acceleration $\boldsymbol{a}_{ij}$ between the states, based on the potential gradient at each modelled position, and written in terms of relative velocity instead of wavenumber, is

$$\gamma_{ij}^{\mathrm{mm}}(t) = \frac{1}{4\sigma^2} \boldsymbol{r}_{ij}(t) \cdot \boldsymbol{v}_{ij}(t) + \frac{\sigma^2 m^2}{\hbar^2} \boldsymbol{v}_{ij}(t) \cdot \boldsymbol{a}_{ij}(t) + \mathrm{i} \frac{m}{2\hbar} \left(|\boldsymbol{v}_{ij}(t)|^2 + \boldsymbol{r}_{ij}(t) \cdot \boldsymbol{a}_{ij}(t) \right), \tag{7.137}$$

where $\boldsymbol{r}_{ij}(t) = \boldsymbol{r}_i(t) - \boldsymbol{r}_j(t)$ and $\boldsymbol{v}_{ij}(t) = \boldsymbol{v}_i(t) - \boldsymbol{v}_j(t)$ are the relative position and velocity of the two states, and $\boldsymbol{a}_{ij}(t)$ is their relative acceleration:

$$\boldsymbol{a}_{ij}(t) = -\frac{1}{m}\left(\nabla V_i(\boldsymbol{r}_i, t) - \nabla V_j(\boldsymbol{r}_j, t)\right). \tag{7.138}$$

Note this is slightly different to the relative acceleration used in the Markovian decoherence calculation in Section 7.5.3. Here, we actually have (approximate) positions $\boldsymbol{r}_i(t)$ and $\boldsymbol{r}_j(t)$ for both states' wavepackets, and so we can evaluate the potential gradient separately at the two positions rather than having to use the position of the state currently selected by the hidden variable for both.

Equation (7.5.4) is our minimal-memory decoherence rate, and requires as input the centre-of-mass position and velocity of each Gaussian wavepacket as input. Let's now move on to how these trajectories are computed.

[TODO: remove consideration of transitions between auxiliary trajectories? Yes, doing this. Need to update comments and footnotes about this earlier in the text]

From many trajectories, one

As mentioned, this method track only one trajectory per state. However, it continuously considers the spawning of new trajectories, and rather than actually tracking multiple trajectories per state or probabilistically replacing the existing trajectories with the new ones, it simply averages the trajectories together, weighted by their quantum probabilities. At the end of each timestep, positions and velocities (other than those of the main trajectory) are updated according to a weighted sum of their present values with those of the main trajectory:

$$\boldsymbol{r}_{i\neq\eta}(t') \rightarrow Q_{i\eta}(t', t)\,\boldsymbol{r}_\eta(t') + \left(1 - Q_{i\eta}(t', t)\right)\boldsymbol{r}_i(t'), \tag{7.139}$$

$$\boldsymbol{v}_{i\neq\eta}(t') \rightarrow Q_{i\eta}^{\text{time}}(t', t)\boldsymbol{v}_\eta(t') + \tilde{Q}_{i\eta}^{\text{space}}(t', t)(\boldsymbol{v}_\eta(t') + \Delta\boldsymbol{v}_{i\eta}(t'))$$
$$+ \left(1 - Q_{i\eta}^{\text{time}}(t', t) - \tilde{Q}_{i\eta}^{\text{space}}(t', t)\right)\boldsymbol{v}_i(t'), \tag{7.140}$$

where $Q_{i\eta}(t', t) = P_{i\eta}(t', t)/|c_i(t')|^2$ is the fraction of the population of state i at time t' that flowed to it from state η in the interval t to t', with $Q_{i\eta}^{\text{time}}(t', t)$ and $\tilde{Q}_{i\eta}^{\text{space}}(t', t)$ defined similarly in terms of $P_{i\eta}^{\text{time}}$ and $\tilde{P}_{i\eta}^{\text{time}}$ as defined in Section 7.4.3, and where $\Delta\boldsymbol{v}_{i\eta}(t')$ is the required velocity correction for a transition from state η to state i state at time t', as discussed in Section 7.4.3:

$$\Delta\boldsymbol{v}_{i\eta}(\boldsymbol{r}, t') = \left[\text{sgn}(\boldsymbol{v}\cdot\hat{\boldsymbol{d}}_{i\eta})\sqrt{(\boldsymbol{v}\cdot\hat{\boldsymbol{d}}_{i\eta})^2 + \frac{2}{m}\left(V_i(\boldsymbol{r}, t') - V_\eta(\boldsymbol{r}, t')\right)} - (\boldsymbol{v}\cdot\hat{\boldsymbol{d}}_{i\eta}),\right]\hat{\boldsymbol{d}}_{i\eta}, \tag{7.141}$$

where $\boldsymbol{r} = \boldsymbol{r}_\eta(t')$ and $\boldsymbol{v} = \boldsymbol{v}_\eta(t')$ are the position and velocity of the main trajectory at time t', and the adiabatic coupling unit vector $\hat{\boldsymbol{d}}_{i\eta} = \hat{\boldsymbol{d}}_{i\eta}(\boldsymbol{r}_\eta(t'))$ is evaluated at the position along the main trajectory at time t'.

The above scheme is constructed to track a mean trajectory for each state, where 'mean' is defined as a weighted sum of positions and velocities with those of the main trajectory whenever transitions from the latter occur, taking into account velocity jumps and frustrated transitions, with the energy conservation calculation for the velocity jumps based on the potentials at the location of the main trajectory.

7.6 Algorithm

Now we have presented all the pieces from which the two versions of the hidden variables semiclassical method are constructed. One more approximation is yet to be mentioned,

and that is that the projected Hamiltonian $\hat{H}_\eta(t)$ first introduced in Section 7.5.1 is approximated by the value of the full Hamiltonian at the centre-of-mass position of the main trajectory's Gaussian wavepacket, without computing an integral:

$$\langle \chi_i | \hat{H}_\eta(t) | \chi_j \rangle = \langle \chi_i \, \phi_\eta(t) | \hat{H}(t) | \chi_j \, \phi_\eta(t) \rangle \qquad (7.142)$$

$$\approx \langle \chi_i \, \boldsymbol{r}_\eta(t) | \hat{H}(t) | \chi_j \, \boldsymbol{r}_\eta(t) \rangle \qquad (7.143)$$

$$\equiv H_{ij}(\boldsymbol{r}_\eta, t). \qquad (7.144)$$

With this final approximation, there are no longer any Guassian wavepackets to compute integrals over or anything else—all details of the wavepackets have been encapsulated and parametrised by the approximations and analytics in the preceding sections, and we can now focus on the classical dynamics of them atoms' centre of mass motion and the quantum evolution of their internal states.

7.6.1 Markovian hidden-variable semiclassical method

This is the simplest version of my method, using crude Markovian decoherence as discussed in Section 7.5.3, and simulating only one trajectory corresponding to the centre-of-mass motion of the state corresponding to the hidden variable at each moment in time.

7.6.2 State variables

The Markovian hidden-variable semiclassical method

Initial conditions

7.6.3 Mean auxiliary trajectories hidden-variables semiclassical method

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!

7.7 Results

Pretty results go here

7.8 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 then 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.

The recognition that surface hopping algorithms are hidden variable theories may be due to me.

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

7.8.1   Comparison with Monte-Carlo wavefunction method and Stochastic Schrödinger equation

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

[DELETE EVERYTHING AFTER THIS AFTER MAKING SURE IT EXISTS ELSEWHERE SATISFACTORILY]

References

[1] M. Genovese. *Research on hidden variable theories: A review of recent progresses.* Physics Reports **413**, 319 (2005). DOI: https://doi.org/10.1016/j.physrep.2005.03.003. [p 1]

[2] S. Aaronson. *Quantum computing and hidden variables.* Phys. Rev. A **71**, 032325 (2005). DOI: 10.1103/PhysRevA.71.032325. [pp 1, 5, 7, 8, and 12]

[3] J. S. Bell. *On the Einstein Podolsky Rosen paradox.* Physics **1**, 195 (1964). [p 1]

[4] M. Schlosshauer, J. Kofler, and A. Zeilinger. *A Snapshot of Foundational Attitudes Toward Quantum Mechanics.* Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics **44**, 222 (2013). ARXIV: 1301.1069, DOI: 10.1016/j.shpsb.2013.04.004. [p 1]

[5] J. C. Tully. *Molecular dynamics with electronic transitions.* The Journal of Chemical Physics **93**, 1061 (1990). DOI: 10.1063/1.459170. [pp 2, 8, 12, 19, and 24]

[6] J. C. Tully. *Mixed quantum-classical dynamics.* Faraday Discuss. **110**, 407 (1998). DOI: 10.1039/A801824C. [p 2]

[7] J. E. Subotnik, A. Jain, B. Landry, A. Petit, W. Ouyang, and N. Bellonzi. *Understanding the Surface Hopping View of Electronic Transitions and Decoherence.* Annual Review of Physical Chemistry **67**, 387 (2016). ARXIV: https://doi.org/10.1146/annurev-physchem-040215-112245, DOI: 10.1146/annurev-physchem-040215-112245. [pp 2, 8, 12, 13, and 18]

[8] J. C. Tully and R. K. Preston. *Trajectory Surface Hopping Approach to Nonadiabatic Molecular Collisions: The Reaction of H+ with D2.* The Journal of Chemical Physics **55**, 562 (1971). ARXIV: https://doi.org/10.1063/1.1675788, DOI: 10.1063/1.1675788. [p 2]

[9] N. Shenvi, J. E. Subotnik, and W. Yang. *Simultaneous-trajectory surface hopping: A parameter-free algorithm for implementing decoherence in nonadiabatic dynamics.* The Journal of Chemical Physics **134**, 144102 (2011). ARXIV: https://doi.org/10.1063/1.3575588, DOI: 10.1063/1.3575588. [pp 2 and 28]

[10] M. F. Herman. *Nonadiabatic semiclassical scattering. I. Analysis of generalized surface hopping procedures.* The Journal of Chemical Physics **81**, 754 (1984). ARXIV: https://doi.org/10.1063/1.447708, DOI: 10.1063/1.447708. [p 2]

[11] G. Granucci, M. Persico, and A. Zoccante. *Including quantum decoherence in surface hopping*. The Journal of Chemical Physics **133**, 134111 (2010). ARXIV: https://doi.org/10.1063/1.3489004, DOI: 10.1063/1.3489004. [p 2]

[12] G. Granucci and M. Persico. *Critical appraisal of the fewest switches algorithm for surface hopping*. The Journal of Chemical Physics **126**, 134114 (2007). ARXIV: https://doi.org/10.1063/1.2715585, DOI: 10.1063/1.2715585. [pp 2 and 8]

[13] A. White, S. Tretiak, and D. Mozyrsky. *Coupled wave-packets for non-adiabatic molecular dynamics: A generalization of Gaussian wave-packet dynamics to multiple potential energy surfaces*. Chem. Sci. **7**, 4905 (2016). DOI: 10.1039/C6SC01319H. [p 2]

[14] J.-Y. Fang and S. Hammes-Schiffer. *Comparison of surface hopping and mean field approaches for model proton transfer reactions*. The Journal of Chemical Physics **110**, 11166 (1999). ARXIV: https://doi.org/10.1063/1.479058, DOI: 10.1063/1.479058. [p 2]

[15] J. E. Subotnik, W. Ouyang, and B. R. Landry. *Can we derive Tully's surface-hopping algorithm from the semiclassical quantum Liouville equation? Almost, but only with decoherence*. The Journal of Chemical Physics **139**, 214107 (2013). ARXIV: https://doi.org/10.1063/1.4829856, DOI: 10.1063/1.4829856. [p 2]

[16] C. J. Billington, C. J. Watkins, R. P. Anderson, and L. D. Turner. *A Monte Carlo wavefunction method for semiclassical simulations of spin-position entanglement*. arXiv:1502.06674 [physics, physics:quant-ph] (2015). ARXIV: 1502.06674. [pp 2 and 11]

[17] S. Dietrich and I. D. Boyd. *Scalar and Parallel Optimized Implementation of the Direct Simulation Monte Carlo Method*. Journal of Computational Physics **126**, 328 (1996). DOI: https://doi.org/10.1006/jcph.1996.0141. [p 2]

[18] K. Mølmer, Y. Castin, and J. Dalibard. *Monte Carlo wave-function method in quantum optics*. J. Opt. Soc. Am. B **10**, 524 (1993). DOI: 10.1364/JOSAB.10.000524. [pp 3 and 23]

[19] M. B. Plenio and P. L. Knight. *The quantum-jump approach to dissipative dynamics in quantum optics*. Rev. Mod. Phys. **70**, 101 (1998). DOI: 10.1103/RevModPhys.70.101. [pp 3 and 23]

[20] W. Gerlach and O. Stern. *Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld*. Zeitschrift für Physik **9**, 349 (1922). DOI: 10.1007/BF01326983. [p 3]

[21] E. Majorana. *Atomi orientati in campo magnetico variabile*. Il Nuovo Cimento (1924-1942) **9**, 43 (1932). DOI: 10.1007/BF02960953. [p 4]

[22] W. Petrich, M. H. Anderson, J. R. Ensher, and E. A. Cornell. *Stable, Tightly Confining Magnetic Trap for Evaporative Cooling of Neutral Atoms*. Phys. Rev. Lett. **74**, 3352 (1995). DOI: 10.1103/PhysRevLett.74.3352. [p 4]

[23] P. Ehrenfest. *Bemerkung über die angenäherte Gültigkeit der klassischen Mechanik innerhalb der Quantenmechanik*. Zeitschrift für Physik **45**, 455 (1927). DOI: 10.1007/BF01329203. [p 4]

[24] S. Aaronson. *Quantum Computing and Dynamical Quantum Models*. (2002). [p 5]

[25] P. Knight.   *The Sinkhorn–Knopp Algorithm: Convergence and Applica-
      tions.*   SIAM Journal on Matrix Analysis and Applications **30**, 261 (2008).
      DOI: 10.1137/060659624. [p 10]

[26] N. Linial, A. Samorodnitsky, and A. Wigderson. *A Deterministic Strongly Polyno-
      mial Algorithm for Matrix Scaling and Approximate Permanents.* Combinatorica
      20, 545 (2000). DOI: 10.1007/s004930070007. [p 10]

[27] P. A. Knight and D. Ruiz. *A fast algorithm for matrix balancing.* IMA Journal of
      Numerical Analysis **33**, 1029 (2013). DOI: 10.1093/imanum/drs019. [p 11]

[28] S. Aaronson. *Quantum Computing Since Democritus.* Cambridge University Press,
      New York, NY, USA (2013). [p 12]

[29] E. Fabiano, G. Groenhof, and W. Thiel.   *Approximate switching algo-
      rithms for trajectory surface hopping.*   Chemical Physics **351**, 111 (2008).
      DOI: https://doi.org/10.1016/j.chemphys.2008.04.003. [p 12]

[30] J. Schwinger, M. O. Scully, and B. G. Englert.   *Is spin coherence like Humpty-
      Dumpty?* Zeitschrift für Physik D Atoms, Molecules and Clusters **10**, 135 (1988).
      DOI: 10.1007/BF01384847. [p 19]

[31] S. Machluf, Y. Japha, and R. Folman. *Coherent Stern–Gerlach momentum splitting on
      an atom chip.* Nature Communications **4**, 2424 (2013). DOI: 10.1038/ncomms3424.
      [p 19]

[32] D. W. Keith, C. R. Ekstrom, Q. A. Turchette, and D. E. Pritchard. *An interferometer
      for atoms.* Phys. Rev. Lett. **66**, 2693 (1991). DOI: 10.1103/PhysRevLett.66.2693.
      [p 19]

[33] E. M. Rasel, M. K. Oberthaler, H. Batelaan, J. Schmiedmayer, and A. Zeilinger.
      Atom Wave Interferometry with Diffraction Gratings of Light. Phys. Rev. Lett. **75**,
      2633 (1995). DOI: 10.1103/PhysRevLett.75.2633. [p 19]

[34] M. A. Schlosshauer. *Decoherence: And the Quantum-To-Classical Transition.* The
      Frontiers Collection. Springer-Verlag, Berlin Heidelberg (2007). [pp 20 and 23]

[35] W. M. Itano, D. J. Heinzen, J. J. Bollinger, and D. J. Wineland. *Quantum Zeno effect.*
      Phys. Rev. A **41**, 2295 (1990). DOI: 10.1103/PhysRevA.41.2295. [p 22]

[36] B. Misra and E. C. G. Sudarshan. *The Zeno's paradox in quantum theory.* Journal
      of Mathematical Physics **18**, 756 (1977). ARXIV: https://doi.org/10.1063/1.523304,
      DOI: 10.1063/1.523304. [pp 22 and 23]

[37] L. Fonda, G. C. Ghirardi, and A. Rimini. *Decay theory of unstable quantum systems.*
      Reports on Progress in Physics **41**, 587 (1978). [p 23]

[38] H. M. Wiseman. *Quantum trajectories and quantum measurement theory.* Quantum
      and Semiclassical Optics: Journal of the European Optical Society Part B **8**, 205
      (1996). [p 23]

[39] G. C. Hegerfeldt. *The Quantum Jump Approach and Quantum Trajectories.* In F.
      Benatti and R. Floreanini (editors), *Irreversible Quantum Dynamics*, volume 622 of
      Lecture Notes in Physics, Berlin Springer Verlag, pages 233–242 (2003). [p 23]

[40] K. Jacobs and D. A. Steck.   *A straightforward introduction to
      continuous quantum measurement.*   Contemporary Physics **47**,
      279 (2006).   ARXIV: https://doi.org/10.1080/00107510601101934,
      DOI: 10.1080/00107510601101934. [p 23]

[41] B. R. Landry and J. E. Subotnik. *How to recover Marcus theory with fewest switches surface hopping: Add just a touch of decoherence.* The Journal of Chemical Physics **137**, 22A513 (2012). ARXIV: https://doi.org/10.1063/1.4733675, DOI: 10.1063/1.4733675. [p 24]

[42] K. F. Wong and P. J. Rossky. *Solvent-induced electronic decoherence: Configuration dependent dissipative evolution for solvated electron systems.* The Journal of Chemical Physics **116**, 8429 (2002). ARXIV: https://aip.scitation.org/doi/pdf/10.1063/1.1468887, DOI: 10.1063/1.1468887. [p 24]

[43] E. R. Bittner and P. J. Rossky. *Quantum decoherence in mixed quantum-classical systems: Nonadiabatic processes.* The Journal of Chemical Physics **103**, 8130 (1995). ARXIV: https://doi.org/10.1063/1.470177, DOI: 10.1063/1.470177. [p 24]

[44] M. Ben-Nun, J. Quenneville, and T. J. Martínez. *Ab Initio Multiple Spawning: Photochemistry from First Principles Quantum Molecular Dynamics.* The Journal of Physical Chemistry A **104**, 5161 (2000). ARXIV: https://doi.org/10.1021/jp994174i, DOI: 10.1021/jp994174i. [p 28]

[45] M. Ben-Nun and T. J. Martínez. *Nonadiabatic molecular dynamics: Validation of the multiple spawning method for a multidimensional problem.* The Journal of Chemical Physics **108**, 7244 (1998). ARXIV: https://doi.org/10.1063/1.476142, DOI: 10.1063/1.476142. [p 28]

[46] S. Yang, J. D. Coe, B. Kaduk, and T. J. Martínez. *An "optimal" spawning algorithm for adaptive basis set expansion in nonadiabatic dynamics.* The Journal of Chemical Physics **130**, 134113 (2009). ARXIV: https://doi.org/10.1063/1.3103930, DOI: 10.1063/1.3103930. [p 28]

Word count

```
Total
Words in text: 48131
Words in headers: 437
Words outside text (captions, etc.): 7469
Number of headers: 107
Number of floats/tables/figures: 35
Number of math inlines: 1820
Number of math displayed: 304
Files: 8
Subcounts:
  text+headers+captions (#headers/#floats/#inlines/#displayed)
  7422+77+1300 (23/3/559/73) File(s) total: atomic_physics.tex
  50+0+0 (0/0/0/0) File(s) total: front_matter.tex
  11176+108+1502 (25/6/339/90) File(s) total: hidden_variables.tex
  1607+8+98 (3/0/12/0) File(s) total: introduction.tex
  23087+155+3403 (31/15/824/137) File(s) total: numerics.tex
  4+22+0 (5/0/0/0) File(s) total: software.tex
  4001+32+781 (11/7/79/4) File(s) total: velocimetry.tex
  784+35+385 (9/4/7/0) File(s) total: wave_mixing.tex
```