

# Stochastic Models of Quantum Decoherence

A Thesis  
Submitted to the Faculty  
of  
Drexel University  
by  
Sam Kennerly  
in partial fulfillment of the  
requirements for the degree  
of  
Doctor of Philosophy  
June 2013

Copyright © 2013 Sam Kennerly.



This work is licensed under the Creative Commons Attribution 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0/> or send a letter to Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, USA.

## Acknowledgments

This thesis benefited from discussions and correspondence with researchers whose expertise in physics, mathematics, and numerical computation was invaluable.

The author's introduction to fundamentals of qubit theory and experiments was motivated, accelerated, and supported by Robert Gilmore and Roberto Ramos.

The *Allyson's Choice* thought experiment was inspired by discussions of Wheeler's delayed-choice experiment with Robert Gilmore and Allyson O'Brien.

The *Zech's Qubit* thought experiment was inspired by discussions with Zechariah Thraikill regarding magnetic shielding for phase qubit experiments supervised by Roberto Ramos.

The *drunken master equation* and its derivation were designed to follow Larry Shepp's advice to "find a martingale" when calculating expectation values of stochastic processes.

The thought experiments in Chapter 3 and conclusions in Chapter 7 were improved and given broader context by Erich Joos' observations regarding *fake decoherence* and David Wineland's suggestion to investigate experiments involving classical noisy fields.

Appendix B is based on Robert Gilmore's suggestion to represent states of  $N$ -level systems as elements of the Lie algebra  $\mathfrak{su}(N)$ .

The simulations in Chapter 6, MATLAB code in Appendix C, and many of the graphics throughout this thesis were produced with help from Allyson O'Brien.

## Table of Contents

<b>LIST OF TABLES</b>	<b>6</b>
<b>LIST OF FIGURES</b>	<b>7</b>
<b>ABSTRACT</b>	<b>9</b>
<b>0 ASSUMPTIONS AND NOTATION</b>	<b>10</b>
0.1 Assumptions . . . . .	10
0.2 Notation . . . . .	11
<b>1 QUBITS AS GEOMETRIC OBJECTS</b>	<b>14</b>
1.1 Density matrices . . . . .	16
1.2 Pauli and Bloch coordinates . . . . .	17
1.3 Time evolution and spin- $\frac{1}{2}$ parameters . . . . .	23
1.4 Generic qubit units: angels, blinks, and rads . . . . .	26
<b>2 DRUNK MODELS</b>	<b>28</b>
2.1 Pathwise construction of a drunk qubit . . . . .	29
2.2 State tomography and statistical inference . . . . .	30
2.3 Fluctuations and dissipation . . . . .	32
2.4 Mixed states and missing information . . . . .	35
2.5 Thermal equilibrium mixed states . . . . .	40
2.5.1 Negative temperatures . . . . .	41
2.5.2 Binary entropy, logit, and logistic functions . . . . .	42
2.5.3 MaxEnt mean states . . . . .	44
<b>3 RANDOM QUBITS</b>	<b>46</b>
3.1 Allyson's choice . . . . .	47
3.1.1 Bob's intended experiment . . . . .	48
3.1.2 Allyson's randomized experiment . . . . .	50
3.1.3 Information was encrypted, not destroyed . . . . .	53
3.2 Decoherence by 1000 small cuts . . . . .	55
3.2.1 A convoluted model . . . . .	55
3.2.2 Generalization to other qubit designs . . . . .	56
3.3 Zech's qubit . . . . .	60
<b>4 STOCHASTIC QUBITS</b>	<b>66</b>
4.1 The Drunken Master Equation . . . . .	68
4.2 The Second Law of Drunk Dynamics . . . . .	71

4.3	Example: Toy model of dephasing . . . . .	74
4.4	Example: Isotropic diffusion . . . . .	77
4.5	Example: Linear Bloch model . . . . .	79
4.6	Example: Nonlinear surplus energy model . . . . .	82
<b>5</b>	<b>NUMERICAL METHODS</b>	<b>85</b>
5.1	Magnus methods . . . . .	87
5.1.1	ExpEuler and ExpMid methods . . . . .	89
5.1.2	Advanced Magnus methods for linear ODEs . . . . .	90
5.2	Matrix exponentials . . . . .	91
5.3	Strong solvers for SDEs . . . . .	93
5.3.1	The Castell-Gaines strategy . . . . .	94
<b>6</b>	<b>SIMULATIONS VS. EXPERIMENTS</b>	<b>96</b>
6.1	Pulse-controlled qubits . . . . .	97
6.1.1	The rotating-wave approximation . . . . .	98
6.1.2	A geometric view: $\pi$ pulses and spiral operators . . . . .	100
6.2	Simulated Rabi cycles . . . . .	104
6.3	Simulated Ramsey fringes . . . . .	108
6.4	Experimental data . . . . .	111
<b>7</b>	<b>CONCLUSIONS</b>	<b>116</b>
7.1	A loophole in Loschmidt's paradox . . . . .	117
7.1.1	Anthropomorphic entropy . . . . .	118
7.1.2	The shuffle hypothesis . . . . .	119
7.1.3	Fake decoherence . . . . .	121
7.2	Comparison with other theories . . . . .	122
7.2.1	Kossakowski-Lindblad master equations . . . . .	124
7.3	Future research . . . . .	126
<b>A</b>	<b>STOCHASTIC CALCULUS</b>	<b>128</b>
A.1	Itô and Stratonovich-Fisk calculus . . . . .	129
A.2	Centrifugal drift of Itô SDEs . . . . .	133
A.3	Probability density for spherical Brownian motion . . . . .	135
<b>B</b>	<b>BASIS OBSERVABLES</b>	<b>138</b>
B.1	$N$ -level systems . . . . .	138
B.2	Fano coordinates . . . . .	140
<b>C</b>	<b>MATLAB CODE</b>	<b>142</b>
C.1	How to use the MATLAB code . . . . .	142
C.2	Scripts . . . . .	143
	<b>LIST OF REFERENCES</b>	<b>153</b>
	<b>VITA</b>	<b>160</b>

---

## List of Tables

1	Vector and matrix notation. . . . .	12
2	Random variable and stochastic process notation. . . . .	13
3	Index conventions. . . . .	13
1.1	State vector $ \Psi\rangle$ vs. pure density matrix $\hat{\rho}_\Psi$ . . . . .	17
1.2	Density matrices vs. Pauli coordinates . . . . .	18
1.3	Coordinate transformation formulas for qubit pure states. . . . .	20
1.4	Rough estimates of natural angular frequency and coherence times. . . . .	27
2.1	Naming conventions for true, sample, mean, and estimated states . . . . .	30
3.1	Example cipher for coin history $THHT\ THTH$ . . . . .	53
6.1	Color-coded volatilities for simulations. . . . .	106

## List of Figures

1.1	Bloch sphere depiction of qubit pure states. . . . .	15
2.1	Fluctuations of estimated state $\mathbf{R}(t)$ for an ideal qubit with $\mathbf{B} = (0, 0, 1)$ . . .	31
2.2	Distribution of sample mean $X$ after 100 measurements of a state with $x = 0$ . . .	33
2.3	vN entropy (in bits) of the canonical density matrix as a function of $kT/\epsilon$ . . .	41
2.4	Probability of detecting excited energy $+\frac{1}{2}\epsilon$ for a thermal-equilibrium qubit. . .	42
2.5	Binary entropy and logit functions in base-2. . . . .	43
3.1	Mach-Zender interferometer for Bob's intended <i>welcher-weg</i> experiment. . . . .	48
3.2	Bob's intended histogram. . . . .	50
3.3	Bob's sabotaged results. . . . .	51
3.4	Allyson's decrypted results. Top: heads trials only. Bottom: tails trials only. . .	52
3.5	Von Neumann entropy (bits) of mean state as a function of $\varphi$ . . . . .	54
3.6	Action of the operator $\hat{\Phi}\hat{S}_1$ on the excited state $(0, 0, -1)$ . . . . .	58
3.7	Action of the GZA on energy eigenstates. . . . .	59
3.8	$P(\text{excited})$ as a function of stray-field amplitude $b$ and duration $\tau$ . . . . .	62
3.9	$P(\text{excited})$ as a function of maximum stray-field strength $\mathcal{B}$ . . . . .	64
3.10	Monte Carlo simulations of Zech's qubit at time $\omega_0 t = 10\pi$ . . . . .	65
4.1	Monte Carlo simulation of toy dephasing model. . . . .	76
4.2	Monte Carlo simulation of isotropic model. . . . .	78
4.3	Monte Carlo simulation of Bloch model. . . . .	81
4.4	Monte Carlo simulation of surplus energy model. . . . .	84
6.1	Simulated $\pi$ pulse and illustration of a $\pi/2$ pulse. . . . .	103
6.2	Full Rabi cycle with $A = \frac{1}{20}$ , $\omega_1 = \omega_0 = 1$ , and $t \in [0, 40\pi]$ . . . . .	104
6.3	Detuned Rabi cycles with $A = \frac{1}{20}$ , $\omega_1 = 0.999$ , $\omega_0 = 1$ , and $t \in [0, 1000]$ . . . .	104
6.4	Simulated $\pi$ pulses using <code>StochasticLinear</code> . . . . .	105
6.5	Rabi cycle master equation simulations. . . . .	106
6.6	Rabi cycle Monte Carlo simulations. . . . .	107
6.7	Ideal Ramsey trials with $\tau = 20\pi$ , $20.5\pi$ , and $21\pi$ . . . . .	109
6.8	Monte Carlo simulation of Ramsey trials with $\tau = 21\pi$ , $T_1 = 400$ , $T_2 = 160$ . . . .	110

---

6.9	RamseyMaster simulation with $T_1 = 400$ , $T_2 = 160$ . . . . .	110
6.10	Rabi cycles and Ramsey fringes for 2 nuclear spin qubits. . . . .	112
6.11	Rabi cycles for 4 phase qubits. . . . .	113
6.12	Rabi cycles and Ramsey fringes for an atomic qubit. . . . .	114
6.13	Rabi cycles, Ramsey fringes, and spin-echo test for a flux qubit. . . . .	115
A.1	42 simulated sample paths for $Y_t$ as $t$ advances from 0 to 100. . . . .	129
A.2	Left: 10 sample paths (Cyclone). Right: 10 sample paths of (Circle). . . . .	134

## ABSTRACT

Stochastic Models of Quantum Decoherence  
Sam Kennerly

Suppose a single qubit is repeatedly prepared and evolved under imperfectly-controlled conditions. A *drunk model* represents uncontrolled interactions on each experimental trial as random or stochastic terms in the qubit's Hamiltonian operator. Time evolution of states is generated by a stochastic differential equation whose sample paths evolve according to the Schrödinger equation. For models with Gaussian white noise which is independent of the qubit's state, the expectation value of the solution obeys a master equation which is identical to the high-temperature limit of the Bloch equation. Drunk models predict that experimental data can appear consistent with decoherence even if qubit states evolve by unitary transformations. Examples are shown in which reversible evolution appears to cause irreversible information loss. This paradox is resolved by distinguishing between the *true state* of a system and the *estimated state* inferred from an experimental dataset.

---

## 0 ASSUMPTIONS AND NOTATION

### 0.1 Assumptions

The following simplifications will be assumed throughout this thesis. These assumptions sacrifice generality for the sake of clarity, simplicity, and timeliness.

#### Orthodox quantum mechanics

This thesis is not intended to advocate or refute any particular interpretation of quantum mechanics. In particular, it does *not* attempt to explain wavefunction collapse as a form of decoherence. However, one must state some assumptions or one has nothing to calculate. Unless explicitly stated otherwise, assume the axioms stated in von Neumann's *Mathematical Foundations of Quantum Mechanics*.<sup>[1]</sup>

#### Finite-dimensional observables

The Hilbert space on which state vectors are defined is assumed to be isomorphic to  $\mathbb{C}^N$  where  $N$  is a natural number. This assumption ensures that linear operators can be represented by  $N \times N$  complex matrices, and it avoids the need for infinite-dimensional operator theory. For the most part, only  $N = 2$  is considered because  $\mathbb{C}^2$  is the natural Hilbert space for a single qubit. Systems with  $N > 2$  are briefly considered in Appendix B.

#### No relativity

Time evolution of unobserved systems is assumed at all times to obey the Schrödinger equation or its density-matrix equivalent, the Liouville-von Neumann (LvN) equation. Special and general relativity are neglected, as is quantum field theory. Conclusions in this thesis may therefore *suggest*, but not *prove*, hypotheses regarding relativistic systems.

## 0.2 Notation

### General notation

- Definitions are indicated by **bold text** and/or the symbol  $\equiv$ .
- Vague, heuristic, or otherwise ill-defined terms are indicated by “quotes.”
- $\Rightarrow$ ,  $\Leftarrow$ ,  $\Leftrightarrow$  denote *implies*, *is implied by*, and *is true if and only if*, respectively.
- $\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$  are natural numbers, integers, reals, and complex numbers, respectively.
- $\{\}$  denote a set with  $|$  denoting conditionals. For example,  $S = \{2n \mid n \in \mathbb{N}, 2n > 8\}$  is the set of all even natural numbers which are greater than 8.
- $\in$  and  $\notin$  denote *is an element of* and *is not an element of*, respectively.
- If  $f$  is a function, then  $f(t)$  is the number returned by  $f$  given input  $t$ . This distinction avoids ambiguous (though usually harmless) terminology such as “the function  $f(t)$ .”

### Vector and matrix notation

The terms *matrix* and *linear operator* are used interchangeably, as are the terms *vector* and *column*. These abuses of notation are committed frequently in this thesis. To be precise: matrices and columns are boxes of numbers which are the components of an operator or vector *with respect to a basis*. If one chooses a different basis for a vector space, then vectors in that space and operators acting on that space do not change. Their *components* change, as do the matrices and columns which represent them.

The symbol  $\mathbb{R}^N$  conventionally denotes the set of all real  $N$ -tuples. Here  $\mathbb{R}^N$  is also assumed to be a vector space with the usual Euclidean dot product as its inner product. If  $\{a_n\}, \{b_n\}$  are components of  $\mathbf{a}$  and  $\mathbf{b}$  with respect to an orthonormal basis, then:

$$\mathbf{a} \cdot \mathbf{b} \equiv \sum_{n=1}^N a_n b_n \qquad |\mathbf{a}| \equiv \sqrt{\mathbf{a} \cdot \mathbf{a}}$$

Similarly,  $\mathbb{C}^N$  is assumed to be equipped with the complex Euclidean inner product:

$$\langle \psi | \varphi \rangle \equiv \sum_{n=1}^N \psi_n^* \varphi_n \qquad \|\psi\| \equiv \sqrt{\langle \psi | \psi \rangle}$$

Other vector and matrix notation conventions are shown in Table 1.

Table 1: Vector and matrix notation.

object	example	notes
real inner product space	$\mathbb{R}^3$	
vector in $\mathbb{R}^N$	$\mathbf{r}$	bold letter
adjoint vector in $\mathbb{R}^N$	$\mathbf{r}^T$	$T$ indicates transpose
inner product of $\mathbb{R}^N$	$\mathbf{r} \cdot \mathbf{s}$	use real Euclidean inner product
complex inner product space	$\mathbb{C}^2$	
vector in $\mathbb{C}^N$	$ \Psi\rangle$	Dirac ket
adjoint vector in $\mathbb{C}^N$	$\langle\Psi $	Dirac bra
inner product of $\mathbb{C}^N$	$\langle\Psi \varphi\rangle$	use complex Euclidean inner product
linear operator or matrix	$\hat{A}$	hat over letter
multiplicative identity matrix	$\hat{1}$	diagonal elements are 1, off-diagonals are 0
adjoint matrix	$\hat{A}^\dagger$	equals conjugate transpose $(\hat{A}^T)^*$
matrix trace	$\text{Tr}[\hat{A}]$	sum of diagonal elements
matrix determinant	$\text{Det}[\hat{A}]$	product of (algebraic) eigenvalues
matrix rank	$\text{Rank}[\hat{A}]$	number of linearly-independent columns

## Special matrices

Pauli matrices:

$$\hat{\sigma}_x \equiv \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \hat{\sigma}_y \equiv \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \hat{\sigma}_z \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

3D rotation generators (not quaternions!):

$$\hat{I} \equiv \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad \hat{J} \equiv \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \quad \hat{K} \equiv \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Cross product as a linear operator:

$$\text{if } \mathbf{r} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{then } [\mathbf{r} \times] \equiv \begin{bmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{bmatrix} = x\hat{I} + y\hat{J} + z\hat{K}$$

## Random variables and stochastic processes

Table 2: Random variable and stochastic process notation.

object	example	notes
probability of event	$P[\mathcal{E}]$	$p_k$ may be used for discrete distributions
expectation value of $X$	$E[X]$	also called <i>mean</i>
variance of $X$	$\text{Var}[X]$	defined as $E[(X - E[X])^2]$
filtration	$\mathcal{F}$	
conditional expectation	$E[X \mathcal{F}]$	pronounced “mean of $X$ over $\mathcal{F}$ ”
conditional probability	$P[\mathcal{E} \mathcal{F}]$	pronounced “probability of $\mathcal{E}$ given $\mathcal{F}$ ”
expectation over natural filtration	$\bar{\mathbf{r}}$	bar over random variable
stochastic process at time $t$	$\mathbf{r}_t$	$\mathbf{r}(t)$ is reserved for deterministic functions
Itô integral	$\int x_t dW$	$dW$ means “use Itô calculus”
Stratonovich-Fisk integral	$\int x_t \circ dW$	$\circ dW$ means “use SF calculus”

## Index conventions

Table 3: Index conventions.

name	symbol	how many of them?
sample times	$\{t_j\}$	$J$ (or $J + 1$ , if $t_0$ is included)
possible Hamiltonians	$\{\hat{H}_k\}$	$K$ (or infinite)
possible states	$\{\hat{\rho}_k\}$	$K$ (or infinite)
components of state vector $ \Psi\rangle$	$\{\Psi_n\}$	$N$
components of real vector $\mathbf{x}$	$\{x_m\}$	$M \equiv N^2 - 1$
basis observables	$\{\hat{A}_m\}$	$N^2$ ( $\hat{A}_0 = \hat{1}$ always)

## 1 QUBITS AS GEOMETRIC OBJECTS

Readers experienced with mixed states and Bloch vectors may prefer to briefly review the *Pauli coordinates* in Section 1.2 or skip this chapter entirely. Most of the material in this chapter can also be found in [2], in which Fano advocates the following viewpoint:

States with less than maximum information, represented by density matrices... can be easily expressed in terms of the mean values of observables. Identifying a state by means of such physical parameters brings out the operational basis of the theory and helps in forming a mental picture.

This thesis focuses primarily on *qubits*, also known as *two-level systems*. A **qubit** is defined here as a system with two linearly-independent energy eigenstates. The lower- and higher-energy eigenstates are denoted  $|0\rangle$  and  $|1\rangle$ . If both have the same energy eigenvalue, the qubit is called **degenerate**, and any two orthogonal states can play the roles of  $|0\rangle$  and  $|1\rangle$ . (Methods for describing  $N$ -level systems, where  $N \in \mathbb{N}$ , are in Appendix B.)

The ordered set  $\{|0\rangle, |1\rangle\}$  is used as a basis, called the **energy basis**, for a complex vector space. A **state vector** is any linear combination of the form:

$$|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad \alpha, \beta \in \mathbb{C}$$

The usual quantum rules of normalization and phase invariance are assumed:

1. A state vector is **normalized** if and only if  $\langle\Psi|\Psi\rangle = |\alpha|^2 + |\beta|^2 = 1$ . State vectors representing physical states are assumed to be normalized.
2. A **phase factor** is any complex number with magnitude 1. Any phase factor can be written in the form  $e^{i\phi}$  for some real  $\phi \in [0, 2\pi)$ . If  $|\Psi\rangle = e^{i\phi}|\Psi'\rangle$  for some  $\phi \in [0, 2\pi)$ , then  $|\Psi\rangle$  and  $|\Psi'\rangle$  are assumed to represent the same physical state.

State vector components  $\alpha, \beta$  can also be written in polar form:

$$|\Psi\rangle = ae^{i\phi_a}|0\rangle + be^{i\phi_b}|1\rangle$$

Phase invariance implies that multiplying  $|\Psi\rangle$  by  $e^{-i\phi_a}$  produces an equivalent state. Normalization implies that  $b = \sqrt{1 - a^2}$ . Thus for any qubit state vector  $|\Psi\rangle \in \mathbb{C}^2$ , it is possible

to choose an equivalent  $|\Psi'\rangle$  such that the  $|0\rangle$  component is positive real:

$$|\Psi'\rangle = a|0\rangle + e^{i\phi}\sqrt{1-a^2}|1\rangle \quad \phi \equiv \phi_b - \phi_a$$

The **Bloch sphere** is a geometric representation of qubit states which is commonly used in experiments. A qubit state is specified by two real coordinates  $(\theta, \phi)$ :

$$|\Psi\rangle = \cos\left(\frac{1}{2}\theta\right)|0\rangle + e^{i\phi}\sin\left(\frac{1}{2}\theta\right)|1\rangle \quad \theta \equiv 2 \arccos(a)$$

Here  $\arccos$  is defined such that  $\theta \in [0, \pi]$ . The set of all states forms a sphere with  $|0\rangle$  at the North pole  $\theta = 0$  and  $|1\rangle$  at the South pole  $\theta = \pi$ .<sup>1</sup> A visualization of the Bloch sphere is shown in Figure 1.1.<sup>2</sup> Note that the lower-energy state is at the *top* of the sphere.

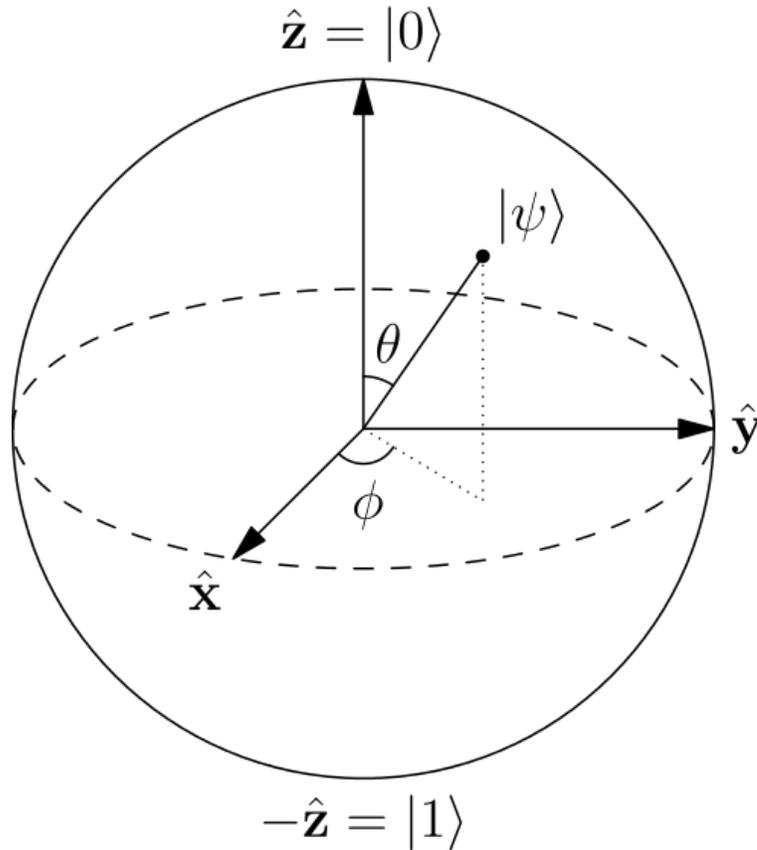


Figure 1.1: Bloch sphere depiction of qubit pure states.

To simplify comparison between theoretical predictions and experimental results, the *Pauli coordinates* used in this thesis are closely related to Bloch-sphere coordinates. Both Pauli and Bloch coordinates can be generalized in a natural way to include *mixed states*.

<sup>1</sup>Bloch-sphere coordinates are singular at the poles; the value of  $\phi$  is undefined there.

<sup>2</sup>Image from Wikimedia Commons by author Glosser.ca. Licensed under Creative Commons BY-SA 3.0.

## 1.1 Density matrices

Instead of  $|\Psi\rangle$ , a state can also be represented by a **pure density matrix**  $\hat{\rho}_\Psi \equiv |\Psi\rangle\langle\Psi|$ :

$$\hat{\rho}_\Psi \equiv |\Psi\rangle\langle\Psi| = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \begin{bmatrix} \alpha^* & \beta^* \end{bmatrix} = \begin{bmatrix} \alpha^*\alpha & \alpha\beta^* \\ \alpha^*\beta & \beta^*\beta \end{bmatrix} \quad (1.1)$$

The symbol  $|\Psi\rangle\langle\Psi|$  is Dirac notation for the outer product of a state vector  $|\Psi\rangle$  with its adjoint vector. If  $\alpha, \beta$  are the components of  $|\Psi\rangle$  with respect to the energy basis, then the components of  $\hat{\rho}_\Psi$  can be calculated as shown in (1.1). A mnemonic for the action of  $|\Psi\rangle\langle\Psi|$  on a vector  $|\varphi\rangle$  is “find the inner product  $\langle\Psi|\varphi\rangle$  and return  $|\Psi\rangle$  times that number.”

For any state vector  $|\Psi\rangle \in \mathbb{C}^2$ , the corresponding density matrix  $\hat{\rho} = |\Psi\rangle\langle\Psi|$  is a rank-1 projection operator. This means  $\hat{\rho}_\Psi$  necessarily has the following properties:

- $\hat{\rho}_\Psi$  is a self-adjoint linear operator:  $\hat{\rho}_\Psi = \hat{\rho}_\Psi^\dagger$ .
- $\hat{\rho}_\Psi$  has trace  $\text{Tr}[\hat{\rho}_\Psi] = 1$  and determinant  $\text{Det}[\hat{\rho}_\Psi] = 0$ .
- $|\Psi\rangle$  is an eigenvector of  $\hat{\rho}_\Psi$  with eigenvalue 1.
- Any vector orthogonal to  $|\Psi\rangle$  is an eigenvector of  $\hat{\rho}_\Psi$  with eigenvalue 0.

These properties remain true for  $N$ -level systems.[2] Pure density matrices provide a formalism for quantum mechanics which is equivalent to the usual state-vector formalism. A major advantage of density matrices is the ability to also represent *mixed states*. An especially useful example from quantum statistical mechanics is the *canonical density matrix* in Section 2.5 which is diagonal with Boltzmann-distributed diagonal elements.

Von Neumann used density matrices to represent systems for which “we do not even know what state is actually present.”[1] If possible state vectors  $\{|\Psi_1\rangle, |\Psi_2\rangle, \dots\}$  are assigned probabilities  $\{p_1, p_2, \dots\}$ , then represent the system with the following density matrix:

$$\bar{\rho} \equiv p_1|\Psi_1\rangle\langle\Psi_1| + p_2|\Psi_2\rangle\langle\Psi_2| + \dots = |\Psi_1\rangle p_1 \langle\Psi_1| + |\Psi_2\rangle p_2 \langle\Psi_2| + \dots$$

Define a **mixed state**, also known as a *statistical mixture*, to be any nontrivial convex combination of pure density matrices.<sup>3</sup> (The pure states need not be orthogonal.) The bar notation  $\bar{\rho}$  instead of  $\hat{\rho}$  is meant to suggest that  $\bar{\rho}$  is an *average* of pure states. A more precise Fano-style interpretation of mixed states is postponed until Section 2.4.

<sup>3</sup>A **convex combination** is a linear combination such that all coefficients are non-negative real and sum to 1. Here *nontrivial* means at least two of the  $\{p_n\}$  are nonzero.

Eigenvalues of a density matrix are non-negative real numbers which sum to 1. A density matrix is **pure** if only one eigenvalue is nonzero; else it is **mixed**. Both pure and mixed density matrices are self-adjoint linear operators, but only pure density matrices are projection operators. Projection operators must satisfy  $\hat{\rho}^2 = \hat{\rho}$ , which is impossible unless eigenvalues of  $\hat{\rho}$  all have magnitude 0 or 1. Two state vectors correspond to the same pure density matrix  $\hat{\rho}$  if and only if they differ by an overall phase factor  $e^{i\varphi}$  for some  $\varphi \in \mathbb{R}$ .

$$|\Psi\rangle = e^{i\varphi}|\Psi'\rangle \quad \Leftrightarrow \quad |\Psi\rangle\langle\Psi| = |\Psi'\rangle\langle\Psi'|$$

If  $\hat{A}$  is a matrix representing an observable, then the same matrix represents the same observable in density-matrix formalism. Suppose  $\{|a_n\rangle\}$  are eigenstates of  $\hat{A}$  with eigenvalues  $\{a_n\}$ . The probability that a measurement returns the result  $a_n$  is:

$$|\langle a_n|\Psi\rangle|^2 = \langle a_n|\Psi\rangle\langle\Psi|a_n\rangle = \langle a_n|\hat{\rho}_\Psi|a_n\rangle$$

If state vectors evolve according to the Schrödinger equation, then density matrices evolve according to the **Liouville-von Neumann (LvN) equation**:

$$i\hbar\frac{d}{dt}\hat{\rho}(t) = [\hat{H}(t), \hat{\rho}(t)]$$

where  $\hat{H}(t)$  is the system's Hamiltonian operator at time  $t$ , and square brackets denote the matrix commutator  $[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$ . Table 1.1 summarizes the distinctions between state-vector and density-matrix formalisms.

Table 1.1: State vector  $|\Psi\rangle$  vs. pure density matrix  $\hat{\rho}_\Psi$ .

object	state vector formalism	density matrix formalism
pure state	$ \Psi\rangle$	$\hat{\rho}_\Psi =  \Psi\rangle\langle\Psi $
probability of result $a_n$	$ \langle a_n \Psi\rangle ^2$	$\langle a_n \hat{\rho}_\Psi a_n\rangle$
expectation value $\langle A \rangle$	$\langle\Psi \hat{A} \Psi\rangle$	$\text{Tr}[\hat{A}\hat{\rho}_\Psi]$
time evolution	$i\hbar\frac{d}{dt} \Psi\rangle = \hat{H} \Psi\rangle$	$i\hbar\frac{d}{dt}\hat{\rho}_\Psi = [\hat{H}, \hat{\rho}_\Psi]$

## 1.2 Pauli and Bloch coordinates

Following Fano's suggestion in [2], any pure or mixed qubit state can be represented by expectation values  $(x, y, z)$  of three observables. A choice common among experimentalists is to use whatever three observables are represented in the energy basis by Pauli matrices.[3] This representation has several practical advantages:

1. 3 real numbers are easier to visualize than a  $2 \times 2$  complex self-adjoint matrix.
2. The vector  $\mathbf{r}$  with Cartesian coordinates  $(x, y, z)$  equals the state's Bloch vector.
3. The time-evolution equation can be written very simply:  $\hbar \dot{\mathbf{r}} = \mathbf{H} \times \mathbf{r}$ .

Table 1.2: Density matrices vs. Pauli coordinates

	density matrices	Pauli coordinates
pure state	$\hat{\rho} =  \Psi\rangle\langle\Psi $	$\mathbf{r} = (x, y, z)$
mixed state	$\bar{\rho} = \sum p_k \hat{\rho}_k$	$\bar{\mathbf{r}} = \sum p_k \mathbf{r}_k$
observable $A$	$2 \times 2$ self-adjoint matrix $\hat{A}$	real $A_w$ and $\mathbb{R}^3$ vector $\mathbf{A}$
possible results $A_{\pm}$	eigenvalues of $\hat{A}$	$A_{\pm} = \frac{1}{2}(A_w \pm  \mathbf{A} )$
eigenstates with results $A_{\pm}$	eigenvectors of $\hat{A}$	$\mathbf{A}_{\pm} = \pm \mathbf{A}/ \mathbf{A} $
probability of result $A_{\pm}$	$\text{Tr}[[A_{\pm}]\langle A_{\pm} \hat{\rho}]$	$\frac{1}{2}(1 + \mathbf{r} \cdot \mathbf{A}_{\pm})$
time evolution	$\hbar \frac{d}{dt} \hat{\rho} = -\imath[\hat{H}, \hat{\rho}]$	$\hbar \dot{\mathbf{r}} = \mathbf{H} \times \mathbf{r}$

Quantum observables are represented by self-adjoint linear operators. For qubit observables, these operators can be represented by  $2 \times 2$  self-adjoint complex matrices. For any qubit observable  $\hat{A}$ , define four real **Pauli coordinates**  $A_w, A_x, A_y, A_z$ :

$$A_w \equiv \text{Tr}[\hat{A}] \quad A_x \equiv \text{Tr}[\hat{A}\hat{\sigma}_x] \quad A_y \equiv \text{Tr}[\hat{A}\hat{\sigma}_y] \quad A_z \equiv \text{Tr}[\hat{A}\hat{\sigma}_z]$$

Any qubit observable is uniquely specified by these four numbers. The inverse map from Pauli coordinates to matrices is:

$$\hat{A} = \frac{1}{2} \begin{bmatrix} A_w + A_z & A_x - \imath A_y \\ A_x + \imath A_y & A_w - A_z \end{bmatrix} = \frac{1}{2} (A_w \hat{1} + A_x \hat{\sigma}_x + A_y \hat{\sigma}_y + A_z \hat{\sigma}_z)$$

Note that any self-adjoint  $2 \times 2$  complex matrix can be written in this form.

The set of all qubit observables forms a real vector space with dimension 4. The set  $\{\hat{1}, \hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\}$  is a basis, and  $A_w, A_x, A_y, A_z$  are the components of  $\hat{A}$  with respect to that basis.<sup>4</sup> The **Hilbert-Schmidt inner product** of two matrices  $\hat{A}, \hat{B}$  is  $\text{Tr}[\hat{A}^\dagger \hat{B}]$ . With this definition, qubit observables are an inner product space and  $\{\hat{1}, \hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\}$  is an orthonormal basis (except for an annoying factor of 2).<sup>5</sup>

$$\text{Tr}[\hat{1}^2] = 2 \quad \text{Tr}[\hat{1}\hat{\sigma}_x] = \text{Tr}[\hat{1}\hat{\sigma}_y] = \text{Tr}[\hat{1}\hat{\sigma}_z] = 0 \quad \text{Tr}[\hat{\sigma}_i \hat{\sigma}_j] = 2\delta_{ij}$$

<sup>4</sup>The set  $\{\hat{1}, \imath\hat{\sigma}_x, \imath\hat{\sigma}_y, \imath\hat{\sigma}_z\}$  is a basis for the quaternion algebra  $\mathbb{H}$ . Pauli coordinates are not quaternions, though quaternions do provide another formalism for qubit calculations.

<sup>5</sup>Many annoying factors of 2 can be removed by using a rescaled inner product  $\text{Tr}[\hat{A}^\dagger \hat{B}]/2$ . But then pure states have  $|\mathbf{r}| = 1/2$ , which conflicts with the usual Bloch-sphere conventions.

For some calculations, it is also convenient to define a real 3D vector  $\mathbf{A}$ :

$$\mathbf{A} \equiv (A_x, A_y, A_z) = \left( \text{Tr}[\hat{A}\hat{\sigma}_x], \text{Tr}[\hat{A}\hat{\sigma}_y], \text{Tr}[\hat{A}\hat{\sigma}_z] \right) \quad |\mathbf{A}| \equiv \sqrt{A_x^2 + A_y^2 + A_z^2}$$

The trace, determinant, and eigenvalues of any qubit observable  $\hat{A}$  are:

$$\text{Tr}[\hat{A}] = A_w \quad \text{Det}[\hat{A}] = \frac{1}{4} \left( A_w^2 - |\mathbf{A}|^2 \right) \quad A_{\pm} = \frac{1}{2} \left( A_w \pm |\mathbf{A}| \right)$$

Density matrices are observables with two extra rules: eigenvalues are non-negative and sum to 1.[2] For any qubit density matrix  $\hat{\rho}$ , define a **Bloch vector**  $\mathbf{r} \in \mathbb{R}^3$ :

$$\mathbf{r} \equiv (x, y, z) \equiv \left( \text{Tr}[\hat{\rho}\hat{\sigma}_x], \text{Tr}[\hat{\rho}\hat{\sigma}_y], \text{Tr}[\hat{\rho}\hat{\sigma}_z] \right)$$

The rule  $\text{Tr}[\hat{\rho}] = 1$  ensures that only three Pauli coordinates  $(x, y, z)$  are needed to uniquely specify  $\hat{\rho}$ . The inverse map from Bloch vectors to density matrices is:

$$\hat{\rho} = \frac{1}{2} \begin{bmatrix} 1+z & x-iy \\ x+iy & 1-z \end{bmatrix} = \frac{1}{2} \left( \hat{1} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z \right)$$

Spherical coordinates of  $\mathbf{r}$  are equivalent to Bloch-sphere coordinates  $(\theta, \phi)$ . To prove this is true, write  $|\Psi\rangle$  as a column vector and write  $\alpha, \beta$  in terms of  $\theta, \phi$ :

$$|\Psi\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \cos(\frac{1}{2}\theta) \\ e^{i\phi} \sin(\frac{1}{2}\theta) \end{bmatrix}$$

Pauli coordinates  $(x, y, z)$  are the expectation values of Pauli matrices  $\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z$ .

$$\begin{aligned} x &= \langle \Psi | \hat{\sigma}_x | \Psi \rangle = \begin{bmatrix} \alpha^* & \beta^* \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = (e^{i\phi} + e^{-i\phi}) \cos(\frac{1}{2}\theta) \sin(\frac{1}{2}\theta) \\ y &= \langle \Psi | \hat{\sigma}_y | \Psi \rangle = \begin{bmatrix} \alpha^* & \beta^* \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = i (e^{-i\phi} - e^{-i\phi}) \cos(\frac{1}{2}\theta) \sin(\frac{1}{2}\theta) \\ z &= \langle \Psi | \hat{\sigma}_z | \Psi \rangle = \begin{bmatrix} \alpha^* & \beta^* \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \cos^2(\frac{1}{2}\theta) - \sin^2(\frac{1}{2}\theta) \end{aligned}$$

These can be simplified with the help of some trigonometric identities:

$$\begin{aligned} e^{i\phi} + e^{-i\phi} &= 2 \cos(\phi) & \cos^2(\tfrac{1}{2}\theta) - \sin^2(\tfrac{1}{2}\theta) &= \cos(\theta) \\ e^{i\phi} - e^{-i\phi} &= 2i \sin(\phi) & \cos(\tfrac{1}{2}\theta) \sin(\tfrac{1}{2}\theta) &= \tfrac{1}{2} \sin(\theta) \end{aligned}$$

In terms of  $\theta$  and  $\phi$ , the Pauli coordinates of any qubit pure state are:

$$x = (\sin \theta)(\cos \phi) \qquad y = (\sin \theta)(\sin \phi) \qquad z = \cos \theta$$

The inverse map from Bloch-sphere to Pauli coordinates is:

$$\theta = \arccos(z) \in [0, \pi] \qquad \phi = \arctan\left(\frac{y}{x}\right) \in [0, 2\pi)$$

The following are thus equally-valid ways to represent a qubit pure state:

**State vector:** a vector  $|\Psi\rangle \in \mathbb{C}^2$  with the usual normalization and phase-factor rules.

**Density matrix:** a rank-1 projection operator  $\hat{\rho}_\Psi \equiv |\Psi\rangle\langle\Psi|$ .

**Bloch coordinates:** two angles  $\theta \in [0, \pi]$  and  $\phi \in [0, 2\pi)$ .

**Pauli coordinates:** a vector  $(x, y, z) \in \mathbb{R}^3$  with  $x^2 + y^2 + z^2 = 1$ .

For reference, Table 1.3 shows pure-state change-of-coordinate formulas.

Table 1.3: Coordinate transformation formulas for qubit pure states.

	$\alpha 0\rangle + \beta 1\rangle$	$(\theta, \phi)$	$(x, y, z)$
$\alpha =$		$\cos(\tfrac{1}{2}\theta)$	$[\tfrac{1}{2}(1+z)]^{1/2}$
$\beta =$		$e^{i\phi} \sin(\tfrac{1}{2}\theta)$	$[\tfrac{1}{2}(1-z)]^{1/2} e^{i \arctan(y/x)}$
$\theta =$	$2 \arccos( \alpha )$		$\arccos(z)$
$\phi =$	$\text{Phase}(\beta) - \text{Phase}(\alpha)$		$\arctan(y/x)$
$x =$	$2\text{Re}[\alpha^*\beta]$	$(\sin \theta)(\cos \phi)$	
$y =$	$2\text{Im}[\alpha^*\beta]$	$(\sin \theta)(\sin \phi)$	
$z =$	$2 \alpha ^2 - 1$	$\cos \theta$	

Pauli and Bloch coordinates can be extended to include mixed states as well. Suppose an unknown pure state might be one of several possible states  $\{|\Psi_1\rangle, \dots, |\Psi_K\rangle\}$  with probabilities  $\{p_1, \dots, p_K\}$ . In Pauli coordinates, von Neumann's definition of mixed state becomes:

$$\bar{\rho} = \sum_{k=1}^K p_k |\Psi_k\rangle\langle\Psi_k| = \sum_{k=1}^K p_k \hat{\rho}_{\Psi_k} = \frac{1}{2} \sum_{k=1}^K p_k (\hat{1} + x_k \hat{\sigma}_x + y_k \hat{\sigma}_y + z_k \hat{\sigma}_z)$$

Define  $\bar{x} \equiv \sum p_k x_k$  to be the expected value of  $x$ , and define  $\bar{y}, \bar{z}$  similarly. Then

$$\bar{\rho} = \frac{1}{2}(\hat{1} + \bar{x}\hat{\sigma}_x + \bar{y}\hat{\sigma}_y + \bar{z}\hat{\sigma}_z)$$

Orthogonality of  $\{\hat{1}, \hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\}$  implies that the Bloch vector of the mixed state  $\bar{\rho}$  is

$$\bar{\mathbf{r}} = \sum_{k=1}^K p_k \mathbf{r}_k = (\bar{x}, \bar{y}, \bar{z})$$

Bloch vectors of mixed states are thus convex combinations of pure-state Bloch vectors. The **convex hull** of a set  $S \subset \mathbb{R}^N$  is the set of all convex combinations of vectors in  $S$ . Define the **Bloch ball** as the convex hull of the Bloch sphere. Points in the Bloch ball can be represented equally well by Pauli coordinates  $(x, y, z)$  with  $x^2 + y^2 + z^2 \leq 1$  or Bloch coordinates  $(r, \theta, \phi)$  with  $r \equiv |\mathbf{r}|$ . The eigenvalues of any qubit density matrix  $\hat{\rho}$  are:

$$\lambda_{\pm} = \frac{1}{2}(1 \pm |\mathbf{r}|)$$

Eigenvalues of  $\hat{\rho}$  must be non-negative, so  $0 \leq |\mathbf{r}| \leq 1$ . Pure density matrices are rank-1 projections with eigenvalues 1 and 0, which requires  $|\mathbf{r}| = 1$ . In other words,

Pure states are *on* the Bloch sphere. Mixed states are *inside* the Bloch sphere.

If two state vectors are orthogonal, then their Bloch vectors point in opposite (not perpendicular!) directions. This apparent contradiction is resolved by distinguishing between different inner products on different spaces. The inner product  $\mathbf{r}_1 \cdot \mathbf{r}_2$  of two Bloch vectors is the usual dot product on  $\mathbb{R}^3$ . The inner product  $\langle \Psi_1 | \Psi_2 \rangle$  of two state vectors is the complex Euclidean inner product on  $\mathbb{C}^2$ . The Hilbert-Schmidt inner product of two matrices  $\hat{A}, \hat{B}$  is  $\text{Tr}[\hat{A}^\dagger \hat{B}]$ . These inner products are related like so:

$$|\langle \Psi_1 | \Psi_2 \rangle|^2 = \text{Tr}[\hat{\rho}_1 \hat{\rho}_2] = \frac{1}{2}(1 + \mathbf{r}_1 \cdot \mathbf{r}_2)$$

For mixed states,  $\langle \Psi_1 | \Psi_2 \rangle$  is not defined, but the relation  $\text{Tr}[\hat{\rho}_1 \hat{\rho}_2] = \frac{1}{2}(1 + \mathbf{r}_1 \cdot \mathbf{r}_2)$  remains valid. The **Bloch angle** between two states is the smallest positive  $\Phi$  such that

$$\cos(\Phi) = \frac{\mathbf{r}_1 \cdot \mathbf{r}_2}{|\mathbf{r}_1| |\mathbf{r}_2|} = \frac{2\text{Tr}[\hat{\rho}_1 \hat{\rho}_2] - 1}{\sqrt{(2\text{Tr}[\hat{\rho}_1^2] - 1)(2\text{Tr}[\hat{\rho}_2^2] - 1)}} \quad (\text{Pure or mixed states})$$

$$\cos(\Phi) = \mathbf{r}_1 \cdot \mathbf{r}_2 = 2\text{Tr}[\hat{\rho}_1 \hat{\rho}_2] - 1 = 2|\langle \Psi_1 | \Psi_2 \rangle|^2 - 1 \quad (\text{Pure states only!})$$

Bloch angles define a metric for pure states: the shortest distance between two points on the Bloch sphere is  $\Phi(\mathbf{r}_1, \mathbf{r}_2) = \arccos(\mathbf{r}_1 \cdot \mathbf{r}_2) \in [0, \pi]$ . Bloch angles must not be confused

with the **quantum angle**  $\Gamma \in [0, \frac{\pi}{2}]$  between two state vectors:<sup>6</sup>

$$\cos(\Gamma) = |\langle \Psi_2 | \Psi_1 \rangle| = \sqrt{\text{Tr}[\hat{\rho}_2 \hat{\rho}_1]} = \sqrt{\frac{1}{2}(1 + \mathbf{r}_1 \cdot \mathbf{r}_2)}$$

For pure states, the Bloch angle  $\Phi$  is *twice* the quantum angle  $\Gamma$ . If two state vectors are orthogonal in  $\mathbb{C}^2$ , then their quantum angle is  $\Gamma = \frac{\pi}{2}$  and their Bloch angle is  $\Phi = \pi$ .

$$\cos(2\Gamma) = 2 \cos^2(\Gamma) - 1 = \mathbf{r}_1 \cdot \mathbf{r}_2 = \cos(\Phi)$$

If one attempts to use this formula to define a quantum angle  $\Gamma$  between two mixed states, then the relation  $\Phi = 2\Gamma$  is no longer valid. The Bloch angle  $\Phi(\mathbf{r}_1, \mathbf{r}_2)$  is a valid metric for any two states with equal magnitudes  $|\mathbf{r}_1| = |\mathbf{r}_2|$ . The quantum angle  $\Gamma(\mathbf{r}_1, \mathbf{r}_2)$  is not. (For example, the quantum angle between  $\mathbf{r} = (0, 0, \frac{1}{2})$  and itself is nonzero.)

The expectation value  $\langle A \rangle$  of an observable  $\hat{A}$  is the inner product  $\text{Tr}[\hat{A}\hat{\rho}]$ . Using orthogonality of  $\{\hat{1}, \hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\}$ , the equivalent statement for Pauli coordinates is:

$$\begin{aligned} \langle A \rangle &\equiv \text{Tr}[\hat{A}\hat{\rho}] = \frac{1}{4} \text{Tr} \left[ (A_w \hat{1} + A_x \hat{\sigma}_x + A_y \hat{\sigma}_y + A_z \hat{\sigma}_z) (\hat{1} + x \hat{\sigma}_x + y \hat{\sigma}_y + z \hat{\sigma}_z) \right] \\ &= \frac{1}{2} (A_w + x A_x + y A_y + z A_z) = \frac{1}{2} (A_w + \mathbf{A} \cdot \mathbf{r}) \end{aligned}$$

Measurement of any qubit observable  $\hat{A}$  has only two possible outcomes: the eigenvalues of  $\hat{A}$ , which are  $A_{\pm} = \frac{1}{2}(A_w \pm |\mathbf{A}|)$ . Call the greater eigenvalue  $A_+$  the “A up” result. If a qubit state is the “A up” eigenstate, then measurement is certain to return the result  $A_+$ . In this case  $\langle A \rangle = A_+$ , which is true if and only if

$$\frac{1}{2}(A_w + \mathbf{A} \cdot \mathbf{r}) = \frac{1}{2}(A_w + |\mathbf{A}|) \quad \Leftrightarrow \quad \mathbf{r} = \frac{\mathbf{A}}{|\mathbf{A}|}$$

In Pauli coordinates, the “A up” eigenstate is represented by a unit vector pointing in the  $\mathbf{A}$  direction. A similar calculation shows that the “A down” eigenstate is represented by a unit vector pointing the opposite direction.

Measurement probabilities are found by projecting  $\mathbf{A}$  onto  $\mathbf{r}$ . Let  $|A_{\pm}\rangle$  be normalized eigenvectors of  $\hat{A}$  with eigenvalues  $A_{\pm}$ . Given a pure state  $|\Psi\rangle$ , these outcomes have probabilities  $P(A_{\pm}) = |\langle A_{\pm} | \Psi \rangle|^2 = \cos^2(\Gamma)$ , where  $\Gamma$  is the quantum angle between  $|A_{\pm}\rangle$  and  $|\Psi\rangle$ .

$$P(A_{\pm}) = \cos^2 \left[ \Gamma(\mathbf{A}_{\pm}, \mathbf{r}) \right] = \cos^2 \left[ \frac{1}{2} \Phi(\mathbf{A}_{\pm}, \mathbf{r}) \right] = \frac{1}{2} \left[ 1 + \cos(\Phi[\mathbf{A}_{\pm}, \mathbf{r}]) \right] = \frac{1}{2} (1 + \mathbf{r} \cdot \mathbf{A}_{\pm})$$

Note that  $A_w = \text{Tr}[\hat{A}]$  does not appear in this formula. Changing the trace of an observable affects the value of its possible results, but not their probabilities. Many calculations can

<sup>6</sup>The quantum angle is also known as the **Fubini-Study metric** on projective 2D complex space  $\mathcal{P}(\mathbb{C}^2)$ .

be simplified by defining the **traceless friend** of a qubit observable  $\hat{A}$ :

$$\hat{A}' \equiv \hat{A} - \frac{1}{2}A_w\hat{1}$$

The trace of  $\hat{A}'$  is  $\text{Tr}[\hat{A}] - \frac{1}{2}A_w\text{Tr}[\hat{1}] = 0$ . An observable and its traceless friend share the same eigenvectors with all eigenvalues shifted by the same additive constant  $\frac{1}{2}A_w$ .

The commutator of any two matrices  $\hat{A}, \hat{B}$  equals the commutator of their traceless friends, which is especially useful for solving the Liouville-von Neumann equation.

$$[\hat{A}', \hat{B}'] = [\hat{A}, \hat{B}] - \frac{1}{2}A_w[\hat{1}, \hat{B}] - \frac{1}{2}B_w[\hat{A}, \hat{1}] + \frac{1}{4}A_wB_w[\hat{1}, \hat{1}] = [\hat{A}, \hat{B}]$$

### 1.3 Time evolution and spin- $\frac{1}{2}$ parameters

Let  $\hat{H}$  be a qubit's (possibly time-dependent) Hamiltonian operator, and assume that time evolution of the qubit's state vector obeys the Schrödinger equation:<sup>7</sup>

$$i\hbar\frac{d}{dt}|\Psi\rangle = \hat{H}|\Psi\rangle$$

The equivalent statement for density matrices is the Liouville-von Neumann (LvN) equation:<sup>[1]</sup>

$$i\hbar\frac{d}{dt}\hat{\rho} = [\hat{H}, \hat{\rho}]$$

Let  $H_w, H_x, H_y, H_z$  be the Pauli coordinates of  $\hat{H}$ . The LvN equation is:

$$\frac{i\hbar}{2}\frac{d}{dt}\left(\hat{1} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z\right) = \frac{1}{4}\left[\left(H_x\hat{\sigma}_x + H_y\hat{\sigma}_y + H_z\hat{\sigma}_z\right), \left(x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z\right)\right]$$

Using bilinearity of the commutator operation, the right-hand side is:

$$\frac{1}{4}\left(H_xy[\hat{\sigma}_x, \hat{\sigma}_y] + H_xz[\hat{\sigma}_x, \hat{\sigma}_z] + H_yx[\hat{\sigma}_y, \hat{\sigma}_z] + H_yz[\hat{\sigma}_y, \hat{\sigma}_z] + H_zx[\hat{\sigma}_z, \hat{\sigma}_x] + H_zy[\hat{\sigma}_z, \hat{\sigma}_y]\right)$$

Every matrix commutes with  $\hat{1}$ , so the coordinate  $H_w = \text{Tr}[\hat{H}]$  does not appear in the LvN equation. Replacing  $\hat{H}$  with its traceless friend  $\hat{H} - \frac{1}{2}\text{Tr}[\hat{H}]\hat{1}$  does not alter time evolution. In physical terms, changing  $\text{Tr}[\hat{H}]$  only shifts the potential energy of the qubit by an arbitrary additive constant. In mathematical terms, changing  $\text{Tr}[\hat{H}]$  changes solutions of the Schrödinger equation only by multiplying  $|\Psi(t)\rangle$  by a meaningless phase factor.

<sup>7</sup>Or at least, it does when nobody is observing the qubit. Drunk models do not attempt to predict or explain evolution of a state during or after measurement.

The Pauli matrix commutators are:

$$[\hat{\sigma}_x, \hat{\sigma}_y] = 2i\hat{\sigma}_z \quad [\hat{\sigma}_y, \hat{\sigma}_z] = 2i\hat{\sigma}_x \quad [\hat{\sigma}_z, \hat{\sigma}_x] = 2i\hat{\sigma}_y$$

The LvN equation then simplifies to:

$$\begin{aligned} \hbar \frac{d}{dt} (x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z) &= (zH_y - yH_z)\hat{\sigma}_x + (xH_z - zH_x)\hat{\sigma}_y + (yH_x - xH_y)\hat{\sigma}_z \\ \frac{d}{dt} \begin{bmatrix} z & x - iy \\ x + iy & -z \end{bmatrix} &= \frac{1}{\hbar} \begin{bmatrix} yH_x - xH_y & zH_y - yH_z - i(xH_z - zH_x) \\ zH_y - yH_z + i(xH_z - zH_x) & -(yH_x - xH_y) \end{bmatrix} \end{aligned}$$

The LvN equation for Pauli coordinates can also be written without complex numbers:

$$\frac{d}{dt} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{\hbar} \begin{bmatrix} 0 & -H_z & H_y \\ H_z & 0 & -H_x \\ -H_y & H_x & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

In terms of 3D rotation generators, the LvN equation is:

$$\hbar \dot{\mathbf{r}} = [H_x \hat{I} + H_y \hat{J} + H_z \hat{K}] \mathbf{r}$$

Here  $\dot{\mathbf{r}}$  is Newton's dot notation for a time derivative. The 3D rotation generators are:

$$\hat{I} \equiv \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad \hat{J} \equiv \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \quad \hat{K} \equiv \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Note that these are *not* quaternions, though quaternions can also be used to represent 3D rotations. The  $\hat{I}, \hat{J}, \hat{K}$  matrices are closely related to the cross product:

$$\mathbf{H} \times \mathbf{r} = [H_x \hat{I} + H_y \hat{J} + H_z \hat{K}] \mathbf{r}$$

A similar calculation shows that, for any qubit observables  $\hat{A}$  and  $\hat{B}$ , the observable  $\hat{C} = -i[\hat{A}, \hat{B}]$  is traceless with Pauli coordinates  $\mathbf{C} = \mathbf{A} \times \mathbf{B}$ .<sup>8</sup>

<sup>8</sup>Hamilton would almost surely recognize this result as *right quaternion multiplication*. Group theorists may recognize it as a Lie algebra isomorphism  $\mathfrak{su}(2) \cong \mathfrak{so}(3)$ . Annoying factors of 2 appear because  $SU(2)$  is a double cover of  $SO(3)$ . The time-evolution equation resembles quaternion multiplication because unit quaternions form the Spin group  $\text{Sp}(1) \cong SU(2)$ .

Using the cross product, the LvN equation can be written in an extremely compact form:

$$\boxed{\hbar \dot{\mathbf{r}} = \mathbf{H} \times \mathbf{r}} \quad (\text{LvN})$$

Equation (LvN) provides a geometric picture of qubit evolution: Bloch vectors rotate about the axis  $\mathbf{H}$  with angular velocity  $\omega = |\mathbf{H}|/\hbar$ . Because  $\mathbf{r} \cdot \dot{\mathbf{r}} = \mathbf{r} \cdot (\mathbf{H} \times \mathbf{r}) = 0$ , the radial velocity of  $\mathbf{r}$  is always zero and states are constrained to a sphere with radius  $|\mathbf{r}(0)|$ .

If  $\mathbf{H}$  is constant in time, then the solution to (LvN) is a matrix exponential:

$$\mathbf{r}(t) = \exp \left[ \frac{1}{\hbar} (H_x \hat{I} + H_y \hat{J} + H_z \hat{K}) t \right] \mathbf{r}(0)$$

**Rodrigues' rotation formula** gives an explicit solution if  $\mathbf{H}$  is constant:[4]

$$\mathbf{r}(t) = \cos(\omega t) \mathbf{r}(0) + \frac{\sin(\omega t) [\mathbf{H} \times \mathbf{r}(0)]}{\hbar \omega} + \frac{[1 - \cos(\omega t)] [\mathbf{H} \cdot \mathbf{r}(0)] \mathbf{H}}{\hbar^2 \omega^2} \quad \omega \equiv \frac{|\mathbf{H}|}{\hbar}$$

If  $\mathbf{H}$  is not constant, then solutions can be approximated using Magnus, Fer, or Dyson series. (Chapter 5 explains a simple Magnus-based numerical method.)

To standardize notation, it is often convenient to refer to a qubit as if it were a spin- $\frac{1}{2}$  particle in a magnetic field  $\mathbf{B}$ . The Hamiltonian for such a system is:

$$\hat{H} = -\gamma \frac{\hbar}{2} (B_x \hat{\sigma}_x + B_y \hat{\sigma}_y + B_z \hat{\sigma}_z)$$

where  $\gamma$  is the particle's gyromagnetic ratio. Pauli coordinates of  $\hat{H}$  are:

$$\mathbf{H} = (H_x, H_y, H_z) = (\text{Tr}[\hat{H} \hat{\sigma}_x], \text{Tr}[\hat{H} \hat{\sigma}_y], \text{Tr}[\hat{H} \hat{\sigma}_z]) = -\gamma \hbar (B_x, B_y, B_z) = -\gamma \hbar \mathbf{B}$$

The LvN equation is then identical in form to the relaxation-free Bloch equation.<sup>9</sup>

$$\hbar \dot{\mathbf{r}} = -\gamma \hbar \mathbf{B} \times \mathbf{r} \quad \Leftrightarrow \quad \dot{\mathbf{r}} = \gamma (\mathbf{r} \times \mathbf{B})$$

To map any physical qubit to a fictional spin- $\frac{1}{2}$  system, do the following:

1. Name the states  $|0\rangle, |1\rangle$  “z-up” and “z-down.” (z-up is the *lower*-energy eigenstate.)
2. Define fictional spin observables  $\{\hat{S}_x, \hat{S}_y, \hat{S}_z\}$  which are  $\frac{1}{2}\hbar$  times the Pauli matrices.
3. Define a vector  $\mathbf{H} = (\text{Tr}[\hat{H} \hat{\sigma}_x], \text{Tr}[\hat{H} \hat{\sigma}_y], \text{Tr}[\hat{H} \hat{\sigma}_z])$ . Define  $\mathbf{B} = -\mathbf{H}$  and  $\gamma = 1/\hbar$ .

This fictional spin provides standardized parameters for different qubit designs.

<sup>9</sup>The LvN and relaxation-free Bloch equations are mathematically equivalent but physically distinct. The LvN equation describes a single qubit. The Bloch equation describes macroscopic nuclear magnetization.

## 1.4 Generic qubit units: angels, blinks, and rads

Qubit designs vary widely: ion traps, optical coherent states, and Josephson junctions are just a few examples.[5][6][7] For a spin- $\frac{1}{2}$  particle, angular momentum is a natural choice for basis observables  $\{\hat{S}_x, \hat{S}_y, \hat{S}_z\}$ . For a charge qubit, Coulombs might more relevant; for an optical qubit, relevant units might be dimensionless photon number or polarization.

Bloch angles and Pauli coordinates are dimensionless. Phrases such as “ $\pi$  pulse” and “ $90^\circ$  rotation” can be interpreted geometrically without reference to the physical design of a particular qubit. But eigenvalues of  $\hat{H}$  are necessarily energies, and the relevant time scales for different qubits can differ by many orders of magnitude. Generic time and energy units can be useful for numerical simulations and comparison of different qubit designs.

For any particular physical qubit, choose some constant reference Hamiltonian  $\hat{H}_0$ . Think of this as “ $\hat{H}$  with all knobs at zero,” i.e. all parameters set to their default values. Let  $E_0, E_1$  denote the energies of states  $|0\rangle$  and  $|1\rangle$ , and define the **energy gap**  $\epsilon \equiv E_1 - E_0$ . The traceless friend of the reference Hamiltonian  $\hat{H}_0$  is then:

$$\hat{H}'_0 = -\frac{1}{2}\epsilon\hat{\sigma}_z = \frac{1}{2} \begin{bmatrix} -\epsilon & 0 \\ 0 & \epsilon \end{bmatrix}$$

Pauli coordinates of  $\hat{H}_0$  are  $\mathbf{H}_0 = (0, 0, -\epsilon)$ . Define the **natural angular frequency** of the qubit to be  $\omega_0 \equiv \epsilon/\hbar$ . The LvN equation and its solution are:

$$\dot{\mathbf{r}} = -\frac{\epsilon}{\hbar}(\mathbf{z} \times \mathbf{r}) = -\omega_0(\mathbf{z} \times \mathbf{r})$$

$$\begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = \begin{bmatrix} \cos(\omega_0 t) & \sin(\omega_0 t) & 0 \\ -\sin(\omega_0 t) & \cos(\omega_0 t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x(0) \\ y(0) \\ z(0) \end{bmatrix}$$

Define new energy and time units such that  $\epsilon, \omega_0$ , and  $\hbar$  can be replaced by 1:

$$\boxed{1 \text{ angel} \equiv \epsilon = E_1 - E_0 \qquad 1 \text{ blink} \equiv \frac{1}{\omega_0} = \frac{\hbar}{E_1 - E_0}}$$

Viewed from above the north pole, the qubit’s Bloch vector  $\mathbf{r}$  rotates clockwise about the  $z$ -axis. A *blink* is the unit of time needed for  $\mathbf{r}$  to rotate 1 radian. The energy unit *angel* is a mnemonic pun on *angular velocity*.<sup>10</sup> In these units,  $\hbar = 1 \text{ angel} \cdot \text{blink}$ .

<sup>10</sup>Fans of British science fiction may also recognize an *angel* as a creature which feeds on potential energy.

Table 1.4 shows order-of-magnitude estimates of timescales for a few qubit designs. (In practice, the time needed to perform a gate operation may be more important than  $\omega_0$ .)

Table 1.4: Rough estimates of natural angular frequency and coherence times.

qubit type	$\omega_0$ (rad/sec)	$T_1, T_2$ (sec)	$T_1, T_2$ (blinks)
superconducting fluxonium [8]	$10^9$	$10^{-6}$	$10^3$
superconducting transmon [8]	$10^9$	$10^{-4}$	$10^5$
atomic dipole trap [9]	$10^9$	$10^{-3}$	$10^6$
diamond nuclear spin [10]	$10^4$	1	$10^4$

Dimensionless energy and time units can also be generalized to higher-dimensional systems. Let  $E_{\max}, E_{\min}$  be the maximum and minimum eigenvalues of  $\hat{H}$  over a given time interval. Define 1 angel to be a unit of energy equal to the *largest* energy gap  $\epsilon \equiv E_{\max} - E_{\min}$ , and define 1 blink to be  $\hbar/\epsilon$ . These generic units can be useful for numerical simulations. For any Hamiltonian  $\hat{H}$ , define a positive-semidefinite  $\hat{H}_+ \equiv \hat{H} - E_{\min}\hat{1}$ . The largest eigenvalue of  $\hat{H}_+$  during the time interval is then 1 angel, and the fastest angular frequency generated by  $\exp[-\frac{i}{\hbar}\hat{H}_+]$  is 1 rad/blink. The Nyquist rate for this frequency is:

$$\frac{2 \text{ samples}}{\text{cycle}} \cdot \frac{\text{cycle}}{2\pi \text{ rad}} \cdot \frac{1 \text{ rad}}{\text{blink}} = \frac{1 \text{ sample}}{\pi \text{ blinks}}$$

If the sample rate of a simulation is faster than 1 sample per  $\pi$  blinks, then it is protected from Nyquist aliasing. Blinks can also be interpreted in terms of Taylor-series convergence. If  $\hat{H}$  is constant in time, then time evolution can be found by a matrix exponential which is defined by a Taylor series:

$$|\Psi(t)\rangle = \exp\left[-\frac{i}{\hbar}\hat{H}t\right]|\Psi(0)\rangle \quad \exp\left[-\frac{i}{\hbar}\hat{H}t\right] = \sum_{m=0}^{\infty} \frac{1}{m!} \left(-\frac{i}{\hbar}\hat{H}t\right)^m$$

Replace  $\hat{H}$  with  $\hat{H}_+ \equiv \hat{H} - E_{\min}\hat{1}$  as before. The **spectral norm**  $\|\hat{M}\|$  of a matrix is the magnitude of its largest eigenvalue, so  $\|\hat{H}_+\| = E_{\max} - E_{\min} = 1$  angel. The spectral norm is **submultiplicative**:  $\|\hat{M}^m\| \leq \|\hat{M}\|^m$ . Thus if  $t \leq 1$  blink, then each Taylor term has smaller spectral norm than the previous term.

A similar result holds for Magnus series solutions to the LvN equation, which are explained in Chapter 5. A sufficient condition for Magnus-series convergence is:

$$\pi > \int_{t_0}^{t_1} \|\hat{H}(t)\| dt$$

The Nyquist condition of  $> 1$  sample per  $\pi$  blinks thus guarantees convergence.

## 2 DRUNK MODELS

Von Neumann used density matrices to represent systems for which “we do not even know what state is actually present.” [1] Drunk models represent experiments for which we do not even know what Hamiltonian operator is actually present during a system’s evolution. An unknown state is represented by a pure-density-matrix-valued stochastic process  $\hat{\rho}_t$ . The **mean state**  $\bar{\rho}(t) \equiv E[\hat{\rho}_t]$  is the expectation value of that process at time  $t$ .<sup>1</sup> In addition to Copenhagen quantum mechanics, drunk models assume the **principle of finite precision**: *Absolutely precise control of an experiment is impossible.*

In principle, the best possible representation of a system is assumed to be some pure state  $\hat{\rho}_t$  and deterministic Hamiltonian  $\hat{H}_t$ . Time evolution of an unobserved system is assumed to obey the Liouville-von Neumann equation exactly. In practice, uncontrolled interactions alter the precise value of  $\hat{H}_t$ , thus altering future values of  $\hat{\rho}_t$ . Given imperfect knowledge of  $\hat{\rho}_t$  and  $\hat{H}_t$ , a scientist cannot precisely predict future states. Instead, he or she can assign probabilities to possible states and Hamiltonians, calculate a mean state, and use that mean state to predict measurement probabilities. Any experiment involving repeated trials then becomes a *de facto* Monte Carlo simulation.

The measurement problem prevents experimenters from uniquely determining the state of a system by performing a single experimental trial.<sup>2</sup> *State tomography* provides an alternative method: perform many measurements of each basis observable, record the results, and infer expectation values from the recorded data. To avoid endless statistical paradoxes, it is helpful to draw sharp distinctions between *true states* of physical systems, *mean states* of theoretical models, and *estimated states* inferred from experimental data. In an ideal experiment, these three quantities converge. In an imperfect experiment, they do not.

Drunk models are not intended to explain wavefunction collapse, nor to resolve the measurement problem. No attempt is made to answer the question, “What happens to a system after a measurement has been made?” The goal is more modest: to show that evolution by random unitary transformations can *appear* to produce von Neumann entropy.

<sup>1</sup>The bar notation  $\bar{\rho}$  is a mnemonic to suggest that  $\bar{\rho}(t)$  is an *average* of possible pure states.

<sup>2</sup>Unorthodox interpretations of quantum mechanics may disagree. Also, *weak measurements* and *non-demolition measurements* are ignored in this thesis for the sake of simplicity. Readers who wish to remedy this omission are encouraged to cite-search [11] and the resulting discussions and controversies.

## 2.1 Pathwise construction of a drunk qubit

Pauli or Bloch coordinates from Chapter 1 allow any qubit density matrix to be represented by a real vector  $\mathbf{r} \in \mathbb{R}^3$  with  $0 \leq |\mathbf{r}| \leq 1$ . The trace of a qubit's Hamiltonian  $\hat{H}$  does not affect time evolution, and its traceless friend  $\hat{H} - \frac{1}{2}\text{Tr}[\hat{H}]$  can be represented by a vector  $\mathbf{H} \in \mathbb{R}^3$ . To ensure compatibility with experimental conventions and for ease of visualization, these vectors are often used instead of  $2 \times 2$  complex matrices throughout this thesis. (Appendix B considers generalizations to  $N$ -level systems.)

Define a **sober qubit** as an initial pure state  $\mathbf{r}_0$  with  $|\mathbf{r}_0| = 1$  and a Hamiltonian  $\mathbf{H}$  which is a deterministic function of time. A **drunk qubit** is an initial pure-state-valued random variable  $\mathbf{r}_0$  and a Hamiltonian  $\mathbf{H}$  which is a stochastic process. (The mathematical formalism of stochastic processes is introduced in Appendix A and its cited references.)

The **pathwise construction** of a drunk qubit is the following algorithm:

1. Define a stochastic process  $\mathbf{H}$  such that each sample path of  $\mathbf{H}$  represents a possible Hamiltonian which might occur on an experimental trial.
2. For each possible Hamiltonian  $\mathbf{H}^k$ , define a **possible state**  $\mathbf{r}_t^k$  as the corresponding solution to the Liouville-von Neumann equation at time  $t$ , given initial condition  $\mathbf{r}_0$ . (The superscript  $^k$  is an index, not an exponent.)
3. Let  $p_k$  denote the probability that sample path  $\mathbf{H}^k$  occurs on a trial. For any  $t \geq 0$ , define a random variable  $\mathbf{r}_t$  by assigning probability  $p_k$  to the  $k$ th possible state  $\mathbf{r}_t^k$ . Call this random variable  $\mathbf{r}_t$  the **true state**.
4. Define the **mean state**  $\bar{\mathbf{r}}(t)$  to be the expectation value  $E[\mathbf{r}_t]$  over all possible states.

If countably-many possible states  $\{\mathbf{r}_t^k\}$  are assigned probabilities  $\{p_k\}$ , then the mean state  $\bar{\mathbf{r}}(t)$  is constructed exactly as in von Neumann's definition:

$$\bar{\mathbf{r}}(t) \equiv E[\mathbf{r}_t] = \sum_k p_k \mathbf{r}_t^k$$

For some of the random models in Chapter 3 and all of the stochastic models in Chapter 4,  $\mathbf{H}$  has uncountably-many sample paths and this pathwise construction becomes inadequate. It is presented here to suggest a physical interpretation: the true state of the system during an experimental trial is some pure-but-unknown  $\mathbf{r}_t$ . Any one of the possible states  $\{\mathbf{r}_t^k\}$  *might* be the true state on a given trial. The mean state  $\bar{\mathbf{r}}(t) = E[\mathbf{r}_t]$  is an average of possible states weighted by their probabilities. The last piece of the puzzle is the *estimated state*  $\mathbf{R}(t)$ , which is often what is depicted in the plots of an experimental research paper. For reference, Table 2.1 summarizes the distinctions.

Table 2.1: Naming conventions for true, sample, mean, and estimated states

Name	Notation	What is it?
True state	$\mathbf{r}_t$	Random variable representing the qubit's unknown state
Possible state	$\mathbf{r}_t^k$	Possible solution to the LvN equation
Mean state	$\bar{\mathbf{r}}(t)$	Expected value $E[\mathbf{r}_t]$ calculated from a drunk model
Estimated state	$\mathbf{R}(t)$	Statistical estimator inferred from experimental data

## 2.2 State tomography and statistical inference

Suppose a scientist wishes to experimentally determine the Pauli coordinates of a qubit. Suppose this scientist measures the observable  $\hat{\sigma}_x$  and the result is  $+1$ . According to orthodox quantum mechanics, the qubit's state is now  $\mathbf{r} = (1, 0, 0)$ . As for what the state *was* before measurement, all that can be said with certainty is  $x \neq -1$ .

**State tomography** is a method for experimentally inferring the state of a system by preparing many identical copies of the system (or by repeatedly evolving the same system under identical conditions) and measuring basis observables – in this case, Pauli matrices.<sup>3</sup> The following procedure can be used to infer the state's Pauli coordinate  $x$ :

1. Prepare the qubit in some state  $\mathbf{r}$ . Measure  $\hat{\sigma}_x$  and record the result.
2. Repeat Step 1 many times using identical preparation procedures for each trial.
3. Find the mean value of the results. Call that number the **sample mean**  $X$ .

If this algorithm is repeated for  $\hat{\sigma}_y$  and  $\hat{\sigma}_z$ , then the resulting data can be used to infer the system's pre-measurement state *ex post facto*. If  $X, Y, Z$  are the sample means of  $\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z$ , then define the **estimated state** like so:

$$\mathbf{R} \equiv (X, Y, Z)$$

Time evolution can be inferred by repeated state tomography. An experimenter chooses  $J$  sample times  $\{t_1, \dots, t_J\}$  and prepares the qubit in some initial state  $\mathbf{r}_0$ . The qubit is evolved for  $t_1$  seconds, then  $\hat{\sigma}_x$  is measured. Many such trials are repeated for each Pauli matrix. The entire state tomography procedure is then repeated with the qubit evolved for  $t_2$  seconds on each trial, and so on until  $J$  estimated states  $\{\mathbf{R}(t_j)\}$  have been calculated. Figure 2.1 shows numerical simulations of a state tomography experiment with finitely many identical, independent measurements of  $\hat{\sigma}_x, \hat{\sigma}_y$ , and  $\hat{\sigma}_z$ .

<sup>3</sup>Any three Hilbert-Schmidt-orthogonal observables can be used instead of Pauli matrices.  $N$ -level basis observables are explained in Appendix B. Non-orthogonal basis observables can also be used; see [12].

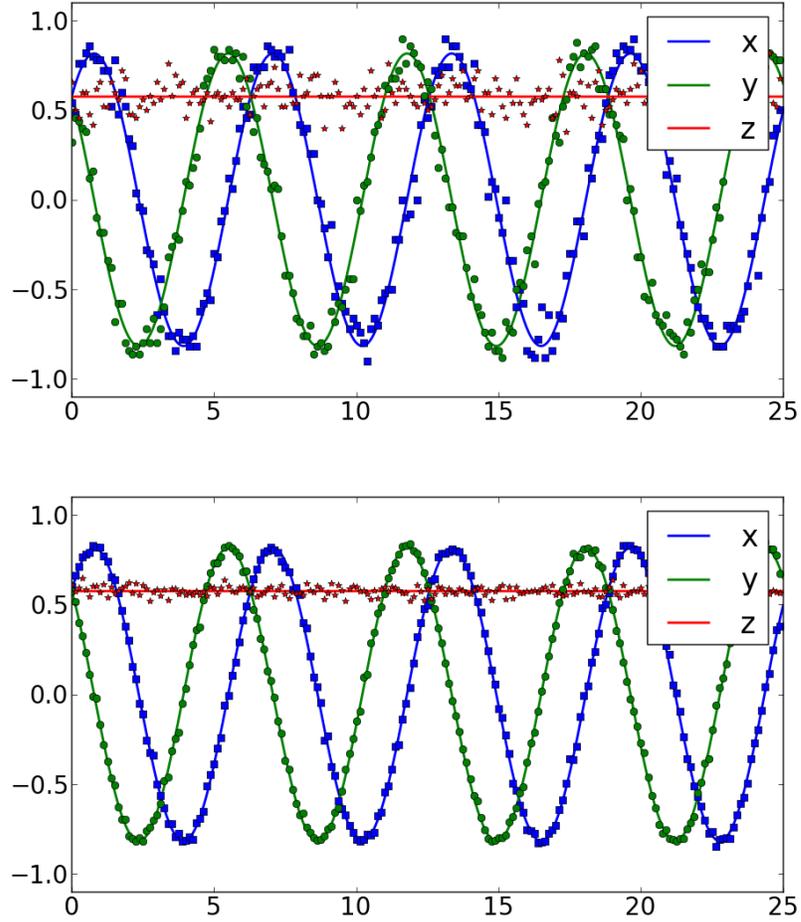


Figure 2.1: Fluctuations of estimated state  $\mathbf{R}(t)$  for an ideal qubit with  $\mathbf{B} = (0, 0, 1)$ . Initial state is  $\mathbf{r}_0 = (1, 1, 1)/\sqrt{3}$ . Solid lines are values of  $x(t), y(t), z(t)$  calculated from the LvN equation. Points are sample means  $X(t_j), Y(t_j), Z(t_j)$  at 200 sample times  $\{t_j\}$ . Top: 100 trials per data point. Bottom: 1000 trials per data point.

The name *estimated state* is meant to emphasize that  $\mathbf{R}$  is not a physical state; it is a *statistical estimator* of a physical state. (More precisely, it is a *maximum-likelihood estimator*.) The distinctions between true, mean, and estimated states may seem unnecessarily complicated, but these three vectors are not subject to the same constraints! True states are assumed to be pure, which constrains  $\mathbf{r}$  to the unit sphere  $|\mathbf{r}| = 1$ . Mean states are convex combinations of true states, which constrains  $\bar{\mathbf{r}}$  to the unit ball  $|\bar{\mathbf{r}}| \leq 1$ . Estimated states are constrained to a cube  $\mathbf{R} \in [-1, 1] \times [-1, 1] \times [-1, 1]$ .

As an extreme example, suppose each Pauli observable is measured only once. Then each of the sample means  $X, Y, Z$  is certain to be either 1 or -1, and the estimated state  $\mathbf{R}$  has square magnitude  $|\mathbf{R}|^2 = 3$ . As a more realistic example, assume the true state is the ground state  $\mathbf{r} = (0, 0, 1)$  and each observable is measured 5000 times. All  $\hat{\sigma}_z$  results are certain to

be 1, so  $Z = 1$ . The probability that all  $\hat{\sigma}_x$  measurements yield the result  $+1$  is  $0.5^{5000}$ . The probability that  $|X| = 1$  is then  $0.5^{4999}$ , as is the probability that  $|Y| = 1$ . So the probability that an estimated state has square magnitude  $|\mathbf{R}|^2 = 3$  is  $0.5^{9998} \approx 2 \cdot 10^{-3010}$ . Estimated states this far “outside the ball” are wildly improbable, but not impossible.

The inference  $\mathbf{r} \approx \mathbf{R}$  is justified by Borel’s version of the Law of Large Numbers (LLN):

If an experiment is repeated a large number of times, independently under identical conditions, then the proportion of times that any specified event occurs approximately equals the probability of the event’s occurrence on any particular trial. The larger the number of repetitions, the better the approximation tends to be.[13]

If  $\mathbf{r}$  is indeed the true state of the system on each trial, then the mean value  $\mathbf{R}$  of infinitely many independent, identical trials equals its expectation value  $\mathbf{r}$ . But there are at least three dangerous words in the preceding sentence: *independent*, *infinite*, and *identical*.

Independence of measurement probabilities follows from a strict reading of the Born rule. If  $\mathbf{r}$  is the qubit’s state, then the probability that a  $\hat{\sigma}_x$  measurement returns the result  $+1$  is completely determined by projecting  $\mathbf{r}$  onto the “ $x$ -up” eigenstate  $\mathbf{X}_+$ :

$$\begin{aligned} P(\mathbf{X}_+|\mathbf{r}) &\equiv P(\hat{\sigma}_x \text{ is measured “up”}|\mathbf{r}) = \frac{1}{2}(1 + \mathbf{r} \cdot \mathbf{X}_+) = \frac{1}{2}(1 + x) \\ P(\mathbf{X}_-|\mathbf{r}) &\equiv P(\hat{\sigma}_x \text{ is measured “down”}|\mathbf{r}) = \frac{1}{2}(1 + \mathbf{r} \cdot \mathbf{X}_-) = \frac{1}{2}(1 - x) \end{aligned}$$

If the Born rule and Copenhagen quantum theory are correct, then the numbers  $P(\mathbf{X}_\pm|\mathbf{r})$  are independent of whatever other procedures scientists might perform. Drunk models blame fluctuation and dissipation of  $\mathbf{R}$  on failure of the *infinite* and *identical* assumptions.

## 2.3 Fluctuations and dissipation

The LvN equation predicts deterministic, well-behaved solutions for  $\mathbf{r}_t$ . Actual experiments produce erratic estimated states  $\mathbf{R}(t)$  which do not exactly follow the LvN equation, even in the unrealistic limit that the Hamiltonian is perfectly controlled and all trials are absolutely identical. Small errors in  $\mathbf{R}(t)$  are practically inevitable because no scientist has enough graduate students to perform  $\infty$  trials of an experiment.

Suppose a qubit is prepared in some true state  $\mathbf{r}$  and a measurement of  $\hat{\sigma}_x$  is performed. The expected value of the sample mean  $X$  after a single measurement of  $\hat{\sigma}_x$  is:

$$E[X|\mathbf{r}] = P(\mathbf{X}_+|\mathbf{r}) - P(\mathbf{X}_-|\mathbf{r}) = \frac{1}{2}(1 + x) - \frac{1}{2}(1 - x) = x$$

The variance of  $X$  after a single trial is only zero if  $x = \pm 1$ :

$$E[(X - x)^2|\mathbf{r}] = E[X^2|\mathbf{r}] - x^2 = 1 - x^2$$

Suppose  $K$  trials are performed and the true state is exactly identical on every trial. This experiment is a sequence of  $K$  independent, identically-distributed (i.i.d.) Bernoulli trials with outcomes  $\pm 1$ . An explicit formula for  $P(X|\mathbf{r})$  is rather awkward-looking, but a sensible derivation can be found by considering the experiment as a sequence of biased coin tosses. Call the measurement result  $+1$  “heads” and the result  $-1$  “tails.” The probability  $P_K(h|\mathbf{r})$  of  $h$  heads results after  $K$  Bernoulli trials is binomially-distributed:

$$P(h|\mathbf{r}) = \frac{K!}{h!(K-h)!} P(\mathbf{X}_+|\mathbf{r})^h P(\mathbf{X}_-|\mathbf{r})^{K-h} = \frac{K!}{2^K h!(K-h)!} (1+x)^h (1-x)^{K-h}$$

The sample mean  $X$  is (number of heads - number of tails) / (number of trials):

$$X = \frac{h - (K - h)}{K} = \frac{2h}{K} - 1 \quad \Leftrightarrow \quad h = \frac{1}{2}K(1 + X)$$

The possible values of  $X$  are:

$$\frac{2h}{K} - 1 \quad h \in \{0, 1, 2, \dots, K\}$$

The expectation value of  $h$  heads results is  $K$  times the probability of heads on a single trial:  $E[h] = KP(\mathbf{X}_+|\mathbf{r}) = \frac{1}{2}K(1+x)$ . The variance of  $h$  is  $K$  times the probability of heads times the probability of tails:  $\sigma_h^2 = KP(\mathbf{X}_+|\mathbf{r})P(\mathbf{X}_-|\mathbf{r}) = \frac{1}{4}K(1-x^2)$ . Using the relation  $X = 2h/K - 1$ , the expected value and variance of  $X$  are:

$$E[X] = \frac{2E[h]}{K} - 1 = x \quad \sigma_X^2 = \left(\frac{2}{K}\right)^2 \sigma_h^2 = \frac{1-x^2}{K}$$

Figure 2.2 shows the probability mass function for  $X$  if  $\hat{\sigma}_x$  is measured for 100 identical states with  $x = 0$ . If many trials are performed, the distribution can be approximated as a normal distribution with mean  $x$  and variance  $(1-x^2)/K$ .

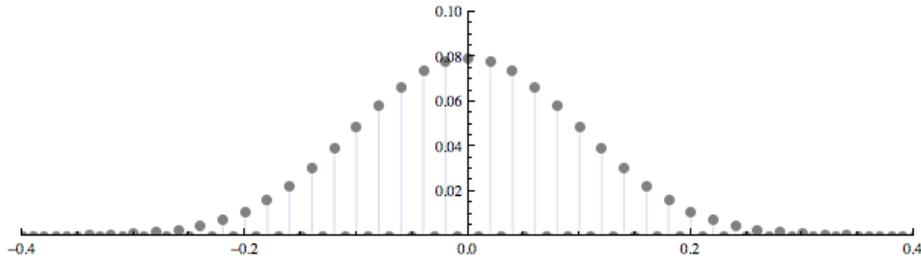


Figure 2.2: Distribution of sample mean  $X$  after 100 measurements of a state with  $x = 0$ .

After finitely-many trials, the observed value of  $\mathbf{R}$  need not exactly equal its expected value. Even if  $\mathbf{r}$  is precisely identical on all  $K$  trials, fluctuations of order  $\sigma_X = \sqrt{(1-x^2)/K}$  will routinely be detected in the sample mean  $X$ , and similarly for  $Y$  and  $Z$ . Experimenters are well aware of the obvious method for reducing fluctuations: do more trials.

The word *identical* is the remaining obstacle to safe use of the Law of Large Numbers. Minimizing fluctuations requires many trials, so experimenters are faced with the challenge of ensuring that many trials are very nearly identical. On any given trial, suppose the qubit's true state  $\mathbf{r}$  might be one of  $K$  possible states  $\{\mathbf{r}^k\}$  with probabilities  $\{p_k\}$ . The probability of an “ $x$ -up” result is given by the **total probability formula**:

$$P(\mathbf{X}_+) = \sum_{k=1}^K P(\mathbf{X}_+|\mathbf{r}^k)P(\mathbf{r}^k) = \sum_{k=1}^K \frac{1}{2} \left(1 + \mathbf{r}^k \cdot \mathbf{X}_+\right) p_k = \frac{1}{2} \left(1 + \bar{\mathbf{r}} \cdot \mathbf{X}_+\right)$$

The term *total probability* refers to the Law of Total Probability. Suppose  $\{D_1, D_2, \dots\}$  is a countable set of pairwise-disjoint events, and suppose  $\sum_k P(D_k) = 1$ . (Less formally: exactly one, but never more than one, of the events  $\{D_k\}$  is certain to occur.) Suppose  $C$  is an event in the same probability space. The discrete Law of Total Probability is:

$$P(C) = \sum_k P(C|D_k)P(D_k)$$

In the context of drunk models, the meaning of this law can be clarified by distinguishing between two different numbers which are both called “probabilities.”

0.  $P(\mathbf{X}_+|\mathbf{r}^k)$  is the probability of an  $x$ -up result, *assuming that  $\mathbf{r}^k$  is the true state*.
1.  $P(\mathbf{r}^k) = p_k$  is the probability that  $\mathbf{r}^k$  is the true physical state of the qubit.

The conditional probability  $P(\mathbf{X}_+|\mathbf{r}^k)$  is determined by the Born rule, which drunk models assume is a fundamental law of Nature. By contrast, the probability  $P(\mathbf{r}^k)$  is assigned by a scientist in order to represent an imprecisely-known state. The latter probabilities are properties of *models*, not properties of physical systems.

For the stochastic models in Chapter 4, the true state  $\mathbf{r}$  has a continuous distribution. Let  $p(\mathbf{r})$  be its probability density. The total probability formula remains valid:

$$\begin{aligned} P(\mathbf{X}_+) &= \iiint P(\mathbf{X}_+|\mathbf{r})p(\mathbf{r}) d^3\mathbf{r} = \iiint \frac{1}{2} \left(1 + \mathbf{r} \cdot \mathbf{X}_+\right) p(\mathbf{r}) d^3\mathbf{r} \\ &= \frac{1}{2} + \frac{1}{2} \left( \iiint \mathbf{r} p(\mathbf{r}) d^3\mathbf{r} \right) \cdot \mathbf{X}_+ = \frac{1}{2} \left(1 + \bar{\mathbf{r}} \cdot \mathbf{X}_+\right) \end{aligned}$$

Confusion often arises when the term *expectation value* is used without specifying what probability space the expectation is calculated over. Let  $E[\mathbf{X}|\mathbf{r}^k]$  denote the conditional

expectation value of a single  $\hat{\sigma}_x$  measurement, given that  $\mathbf{r}^k$  is the true state of the system. This expectation is a weighted average over both possible measurement results  $\pm 1$ . The Born rule shows that  $E[X|\mathbf{r}^k]$  is just the  $x$ -coordinate of  $\mathbf{r}^k$ :

$$E[X|\mathbf{r}^k] = P(\mathbf{X}_+|\mathbf{r}^k) - P(\mathbf{X}_-|\mathbf{r}^k) = \frac{1}{2}(1 + \mathbf{r}^k \cdot \mathbf{X}_+) - \frac{1}{2}(1 - \mathbf{r}^k \cdot \mathbf{X}_+) = \mathbf{r}^k \cdot \mathbf{X}_+$$

The **total expectation value** of a qubit measurement is an average over both possible measurement results *and all possible states*:

$$E[X] \equiv \sum_k E[X|\mathbf{r}^k]P(\mathbf{r}^k) = \sum_k (\mathbf{r}^k \cdot \mathbf{X}_+)p_k = \bar{\mathbf{r}} \cdot \mathbf{X}_+ = \bar{x}$$

The total expectation values of  $Y$  and  $Z$  are defined similarly. The components of the estimated state  $\mathbf{R}$  are the sample means  $X, Y, Z$ , so the total expectation of  $\mathbf{R}$  is:

$$E[\mathbf{R}] = (E[X], E[Y], E[Z]) = (\bar{x}, \bar{y}, \bar{z}) = \bar{\mathbf{r}}$$

Drunk models do not assume the true state  $\mathbf{r}$  is identical on all trials of an experiment. The true state is represented by a random variable (or stochastic process) which is identically *distributed* on all trials. Given this assumption, measurement probabilities are predicted by the total probability formula, which itself is derived from the Born rule and whatever probability distribution is assigned to  $\mathbf{r}$ . In the limit that infinitely-many not-quite-identical trials are performed, the estimated state  $\mathbf{R}$  converges to the mean state  $\bar{\mathbf{r}}$ .

**Dissipation** is defined here as the tendency of estimated states  $\mathbf{R}(t)$  to drift inward as  $t$  increases. The Liouville-von Neumann equation does not permit true states to leave the Bloch sphere. Mean states are *not* subject to the same constraint; the radius  $|\bar{\mathbf{r}}(t)|$  often changes as  $t$  increases. Pauli and Bloch coordinates provide a geometric picture: possible states stay on the surface of the Bloch sphere, but the average of all possible states wanders around the interior of the Bloch ball. Section 2.4 shows that dissipation can be given a Shannon-inspired interpretation as *loss of information* about a qubit's true state.

## 2.4 Mixed states and missing information

Fano interpreted mixed states as “states of less than maximum information.” [2] Drunk models use a similar interpretation: a mean state is defined out of practical necessity whenever a scientist is unable to precisely determine a system's state vector. The phrase “less than maximum information” is replaced here by “greater than zero von Neumann entropy.”

The **Shannon entropy** of a discrete probability distribution  $\{p_k\}$  is defined:[14]

$$S[\{p_k\}] \equiv - \sum_k p_k \log(p_k) \quad (\text{S Entropy})$$

A simple mnemonic for (S Entropy) is to define the **surprisal** of an event to be  $\log(1/p)$ , where  $p$  is the probability of that event occurring. Shannon entropy is then the “expected surprisal”  $\sum p_k \log(1/p_k)$ . The choice of logarithm base determines units: **bit** for base-2, **nat** for base- $e$ , and **digit** for base-10.<sup>4</sup> These dimensionless units are not physical measurements, but merely reminders of which logarithm was used to calculate  $S$ .

If one or more of the  $\{p_k\}$  are zero, then the expressions  $-\log(p_k)$  and  $\log(1/p)$  diverge. (One would be infinitely surprised to observe an impossible event!) To avoid this dilemma, the definition of Shannon entropy can be modified slightly by replacing the expression  $0 \log(0)$  with  $\lim_{p \rightarrow 0^+} p \log(p) = 0$  so that impossible events do not contribute to the sum.

According to Shannon, von Neumann was responsible for the name *entropy*:

My greatest concern was what to call it. I thought of calling it “information,” but the word was overly used, so I decided to call it “uncertainty.” When I discussed it with John von Neumann, he had a better idea. Von Neumann told me, “You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, nobody knows what entropy really is, so in a debate you will always have the advantage.”[15]

The word *uncertainty*, as used in quantum mechanics, also has a connection to Shannon entropy. The Heisenberg uncertainty principle is well-known, but the **entropic uncertainty principle** is a stronger statement! Suppose that  $\Psi(x)$  is a wavefunction and  $\tilde{\Psi}(k)$  its (unitary) Fourier transform. Then  $|\Psi(x)|^2$  is a probability density for position and  $|\tilde{\Psi}(k)|^2$  is a probability density for momentum. If these functions have well-defined Shannon entropies  $S[|\Psi|^2], S[|\tilde{\Psi}|^2]$ , then the sum of their entropies is bounded from below:

$$S[|\Psi|^2] + S[|\tilde{\Psi}|^2] \geq \log(\pi e)$$

This result was conjectured by Hirschman and Everett and later proved by Beckner.[16][17][18] A proper discussion of continuous Shannon entropy and entropic uncertainty is outside the bounds of this thesis; see [19] and its cited references for a survey of recent research.

<sup>4</sup>If surprisal of an event  $\mathcal{E}$  is 4 Digits, then  $P[\mathcal{E}] = 10^{-4}$  and  $P[\text{not } \mathcal{E}] = 0.9999$  has 4 significant digits.

The **von Neumann entropy** of a density matrix  $\bar{\rho}$  is defined:[1]

$$S[\bar{\rho}] \equiv -\text{Tr}[\bar{\rho} \log(\bar{\rho})] \quad (\text{vN Entropy})$$

To see the relation between (S Entropy) and (vN Entropy), recall that any self-adjoint matrix  $\bar{\rho}$  can be diagonalized  $\bar{\rho} = \hat{U} \hat{D} \hat{U}^{-1}$  where  $\hat{U}$  is unitary and  $\hat{D}$  is diagonal and real. The **matrix logarithm** of a non-negative self-adjoint matrix  $\bar{\rho}$  is:

$$\log(\bar{\rho}) \equiv \hat{U} \log(\hat{D}) \hat{U}^{-1}$$

where  $\log(\hat{D})$  is defined in the obvious way: replace the diagonal elements of  $\hat{D}$  with their natural logarithms. (Matrix exponentials are covered in more detail in Chapter 5.) Then  $\log$  is the inverse function of the matrix exponential:

$$\exp[\log(\bar{\rho})] = \exp[\hat{U} \log(\hat{D}) \hat{U}^{-1}] = \hat{U} \exp[\log(\hat{D})] \hat{U}^{-1} = \hat{U} \hat{D} \hat{U}^{-1} = \bar{\rho}$$

The von Neumann entropy of  $\bar{\rho}$  can also be written:

$$S[\bar{\rho}] \equiv -\text{Tr}[\bar{\rho} \log(\bar{\rho})] = -\text{Tr}[\hat{U} \hat{D} \hat{U}^{-1} \hat{U} \log(\hat{D}) \hat{U}^{-1}] = -\text{Tr}[\hat{U} \hat{D} \log(\hat{D}) \hat{U}^{-1}]$$

The trace of a matrix is invariant under unitary transformations, so

$$S[\bar{\rho}] = -\text{Tr}[\hat{D} \log(\hat{D})] = -\sum_{n=1}^N \lambda_n \log(\lambda_n)$$

where  $\{\lambda_n\}$  are the diagonal elements of  $\hat{D}$ , which are necessarily the eigenvalues of  $\bar{\rho}$ . The eigenvalues  $\{\lambda_n\}$  of any density matrix are non-negative and sum to 1, so the set  $\{\lambda_n\}$  satisfies the axioms of a discrete probability distribution. A mnemonic for this result is: *The von Neumann entropy of a density matrix equals the Shannon entropy of its eigenvalues*

Let  $\{|\varphi_n\rangle\}$  denote the eigenvectors of  $\bar{\rho}$  with eigenvalues  $\{\lambda_n\}$ . It is tempting, but misleading, to interpret each eigenvalue  $\lambda_n$  as “the probability the system is in state  $|\varphi_n\rangle\langle\varphi_n|$ .” The potential confusion is: spectral decomposition of  $\bar{\rho}$  is unique, but two distinct convex combinations of pure states can result in the same  $\bar{\rho}$ . For example, suppose a qubit state might be either  $x$ -up or  $z$ -up, each with probability  $\frac{1}{2}$ . The mean state  $\bar{\rho}$  is:

$$\bar{\rho} = \frac{1}{2}|x_\uparrow\rangle\langle x_\uparrow| + \frac{1}{2}|z_\uparrow\rangle\langle z_\uparrow| = \frac{1}{4} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2} \left( \hat{1} + \frac{1}{2} \hat{\sigma}_x + \frac{1}{2} \hat{\sigma}_z \right)$$

The eigenvalues of  $\bar{\rho}$  are  $\lambda_{\pm} = \frac{1}{2}(1 \pm \frac{1}{\sqrt{2}})$ . Its spectral decomposition is:

$$\bar{\rho} = \lambda_+ |\varphi_+\rangle\langle\varphi_+| + \lambda_- |\varphi_-\rangle\langle\varphi_-| = \frac{1 + \frac{1}{\sqrt{2}}}{4} \begin{bmatrix} 1 + \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 1 - \frac{1}{\sqrt{2}} \end{bmatrix} + \frac{1 - \frac{1}{\sqrt{2}}}{4} \begin{bmatrix} 1 - \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & 1 + \frac{1}{\sqrt{2}} \end{bmatrix}$$

The calculation is easier in terms of Pauli coordinates. Eigenstates of  $\bar{\mathbf{r}}$  are unit vectors in the  $\pm\bar{\mathbf{r}}$  directions with eigenvalues  $\lambda_{\pm} = \frac{1}{2}(1 + |\bar{\mathbf{r}}|)$ . The mean state is:

$$\bar{\mathbf{r}} = \frac{1}{2}(1, 0, 0) + \frac{1}{2}(0, 0, 1) = \frac{1}{2}(1, 0, 1)$$

Its eigenstates  $\mathbf{r}_{\pm}$  and eigenvalues  $\lambda_{\pm}$  are:

$$\mathbf{r}_{\pm} = \pm \frac{1}{\sqrt{2}}(1, 0, 1) \quad \lambda_{\pm} = \frac{1}{2}(1 \pm \frac{1}{\sqrt{2}})$$

The qubit's state was assumed to be either  $(1, 0, 0)$  or  $(0, 0, 1)$ , so it is misleading to say “the qubit state is  $\mathbf{r}_+$  or  $\mathbf{r}_-$  with probabilities  $\lambda_{\pm}$ .” A more accurate statement would be:

Suppose the system's state is a random variable with outcomes  $\{\hat{\rho}_k\}$  and probabilities  $\{p_k\}$ . The mean state is  $\bar{\rho} = \sum p_k \hat{\rho}_k$ . Let  $|\lambda_k\rangle$  be eigenvectors of  $\bar{\rho}$  with eigenvalues  $\lambda_k$ . If we *pretend* that  $\{|\lambda_k\rangle\langle\lambda_k|\}$  are the only possible states, with probabilities  $\lambda_{\pm}$ , then we find the same value for  $\bar{\rho}$ .

This unwieldy statement is probably not worth memorizing. The point is: many different probability distributions for  $\hat{\rho}$  can correspond to the same mixed state  $\bar{\rho}$ .

In terms of Pauli coordinates, the eigenvalues of a qubit mixed state  $\bar{\mathbf{r}}$  are  $\frac{1}{2}(1 \pm |\bar{\mathbf{r}}|)$ . The von Neumann entropy of that mixed state is then:

$$\begin{aligned} S(\bar{\mathbf{r}}) &= -\frac{1 + |\bar{\mathbf{r}}|}{2} \log\left(\frac{1 + |\bar{\mathbf{r}}|}{2}\right) - \frac{1 - |\bar{\mathbf{r}}|}{2} \log\left(\frac{1 - |\bar{\mathbf{r}}|}{2}\right) \\ &= \log(2) - \frac{1}{2} \left[ (1 + |\bar{\mathbf{r}}|) \log(1 + |\bar{\mathbf{r}}|) + (1 - |\bar{\mathbf{r}}|) \log(1 - |\bar{\mathbf{r}}|) \right] \end{aligned}$$

Note that  $S(\bar{\mathbf{r}})$  is a function of  $|\bar{\mathbf{r}}|$  only, and  $S \rightarrow 0$  only in the limit  $|\bar{\mathbf{r}}| \rightarrow 1$ . This is the precise meaning of “pure states are on the Bloch sphere and mixed states are in the Bloch ball.” The center  $\bar{\mathbf{r}} = \mathbf{0}$  is the maximum-entropy mixed state with  $S(\bar{\mathbf{r}}) = \log(2)$ .

Shannon interpreted his formula as a measure of “missing information” in a random variable.<sup>[14]</sup>

Von Neumann entropy of a mean state has a similar interpretation as the “minimum ignorance” for a measurement. If the mean state is  $\bar{\mathbf{r}}$  and an observable  $\mathbf{A}$  is measured, then the total probability of the “heads” result is:

$$P(\mathbf{A}_+) = \frac{1}{2}(1 + \bar{\mathbf{r}} \cdot \mathbf{A}_+)$$

To a probabilist, any qubit measurement is a coin toss: a random variable with two outcomes arbitrarily named “heads” and “tails.” The Shannon entropy of this coin toss is:

$$I(\mathbf{A}, \bar{\mathbf{r}}) = -P(A_+|\bar{\mathbf{r}}) \log[P(A_+|\bar{\mathbf{r}})] - P(A_-|\bar{\mathbf{r}}) \log[P(A_-|\bar{\mathbf{r}})]$$

Note that  $I(\mathbf{A}, \bar{\mathbf{r}})$  is used here to denote Shannon entropy of a single measurement to avoid confusion with  $S(\bar{\mathbf{r}})$ , the von Neumann entropy of a mean state.

If  $|\bar{\mathbf{r}}| = 1$  and  $\hat{A}$  is an observable whose Bloch vector points in the same direction as  $\bar{\mathbf{r}}$ , then measurement of  $\hat{A}$  is certain to return the value  $\frac{1}{2}(A_w + |\mathbf{A}|)$ . But if  $|\bar{\mathbf{r}}| < 1$ , then there is *no* observable whose result is certain.  $\mathbf{A}_+$  is a unit vector by definition (see Section 1.2), so the dot product  $\bar{\mathbf{r}} \cdot \mathbf{A}_+$  is necessarily between  $-|\bar{\mathbf{r}}|$  and  $|\bar{\mathbf{r}}|$ . If  $\bar{\mathbf{r}}$  is “inside the ball,” then the total probability formula must be less than 1 no matter how  $\mathbf{A}$  is chosen.

For example, let  $\bar{\mathbf{r}} = (r, 0, 0)$  for some  $r < 1$ . Then the result of a  $\hat{\sigma}_x$  measurement is 1 with probability  $\frac{1}{2}(1+r)$ , and the measurement entropy is

$$I(\mathbf{A}, \bar{\mathbf{r}}) = -\frac{1+r}{2} \log\left(\frac{1+r}{2}\right) - \frac{1-r}{2} \log\left(\frac{1-r}{2}\right) = S(\bar{\mathbf{r}})$$

For qubits, von Neumann entropy is a lower bound on the Shannon entropy of a single measurement.<sup>5</sup> More precisely,  $I(\mathbf{A}, \bar{\mathbf{r}}) \geq S(\bar{\mathbf{r}})$  with equality only when  $\mathbf{A}$  and  $\bar{\mathbf{r}}$  point the same or opposite directions. For the maximum-entropy mixed state  $|\bar{\mathbf{r}}| = 0$ , the lower bound for measurement entropy of *any* observable is  $S(\bar{\mathbf{r}}) = \log(2) = 1$  bit, and all measurements become fair coin tosses. To paraphrase von Neumann, the statement  $\bar{\mathbf{r}} = \mathbf{0}$  is equivalent to saying “we have *absolutely no idea* what state is actually present.”

Note that the LvN equation conserves  $|\bar{\mathbf{r}}|$  and thus cannot alter  $S(\bar{\mathbf{r}})$ . Even in higher dimensions, the LvN equation evolves density matrices by unitary transformations which cannot change eigenvalues of  $\hat{\rho}$  and thus cannot alter von Neumann entropy. Zurek summarized this result concisely: “Unitary evolution condemns states to purity.” [20]

For all the models in this thesis, possible states are assumed to obey the LvN equation. What gains entropy are the mean state of a model and the estimated state inferred from an experiment. If true, mean, and estimated states are accidentally confused with each other, then qubits *appear* to decohere irreversibly. Drunk models tell a different story: true states evolve reversibly, and scientists simply lose track of them. The thought experiments in Chapter 3 are meant to suggest that any information “missing” from a mean state should be considered *encrypted* rather than *destroyed*.

<sup>5</sup>Caution is required when generalizing this result to higher-dimensional systems. For example, consider orbital angular momentum eigenstates  $\{|l, m\rangle\}$ . The mixture  $\hat{\rho} = \frac{1}{3}|1, 1\rangle\langle 1, 1| + \frac{1}{3}|1, 0\rangle\langle 1, 0| + \frac{1}{3}|1, -1\rangle\langle 1, -1|$  has  $S[\hat{\rho}] = \log(3)$ . Measurement of any component of  $\mathbf{L}$  is equally likely to produce  $\hbar, 0$ , or  $-\hbar$ , but the result of measuring  $\hat{L}^2$  is *certain* to be  $2\hbar^2$ . The Shannon entropy an  $\hat{L}^2$  measurement is zero for this mixture.

## 2.5 Thermal equilibrium mixed states

One method for finding thermal-equilibrium mixed states is to calculate a **canonical density matrix** which is diagonal with Boltzmann-distributed energies.[21] Let  $\{|n\rangle\}$  denote energy eigenstates with eigenvalues  $\{E_n\}$ , and define inverse temperature  $\beta \equiv 1/(kT)$ . (The  $k$  is Boltzmann's constant.) The canonical density matrix is:

$$\bar{\rho} = \frac{1}{Z} \sum_{n=1}^N |n\rangle\langle n| e^{-\beta E_n} \quad Z \equiv \sum_{n=1}^N e^{-\beta E_n}$$

The probability of energy measurement result  $E_n$  is the diagonal element  $\bar{\rho}_{nn}$ , which is assumed to be proportional to its Boltzmann factor  $e^{-\beta E_n}$ . The **partition function**  $Z$  is a normalization factor to ensure  $\text{Tr}[\bar{\rho}] = 1$ . The sums are over all energy *eigenstates*, not over all *eigenvalues*. For degenerate systems, the distinction is important: if multiple orthogonal states share the same eigenvalue  $E_n$ , then the factor  $e^{-\beta E_n}$  is counted multiple times. Note also that adding the same constant to all eigenvalues  $E_n$  does not change  $\bar{\rho}$ .

In Section 2.4, the surprisal of an event with probability  $P$  was defined to be  $\log(1/P) = -\log(P)$ . For nondegenerate systems, assigning a Boltzmann distribution to energy measurement probabilities is equivalent to claiming that a plot of  $\text{Surprisal}(E_n)$  consists of points on a line with slope  $\beta$  and  $y$ -intercept  $\log(Z)$ .<sup>6</sup>

$$\text{Surprisal}(E_n) \equiv -\log(P(E_n)) = -\log\left(\frac{1}{Z}e^{-\beta E_n}\right) = \log(Z) + \beta E_n$$

For a qubit with the usual reference Hamiltonian  $\hat{H} = -\frac{1}{2}\epsilon\hat{\sigma}_z$ , the two possible energies are  $E_0 = -\frac{1}{2}\epsilon$  and  $E_1 = +\frac{1}{2}\epsilon$ . The canonical qubit density matrix is:

$$\bar{\rho} = \frac{|0\rangle\langle 0|e^{\frac{1}{2}\beta\epsilon} + |1\rangle\langle 1|e^{-\frac{1}{2}\beta\epsilon}}{e^{\frac{1}{2}\beta\epsilon} + e^{-\frac{1}{2}\beta\epsilon}} = \frac{1}{2 \cosh(\frac{1}{2}\beta\epsilon)} \begin{bmatrix} e^{\frac{1}{2}\beta\epsilon} & 0 \\ 0 & e^{-\frac{1}{2}\beta\epsilon} \end{bmatrix}$$

The energy measurement probabilities for this density matrix are:

$$P(\text{ground}) = \bar{\rho}_{00} = \frac{e^{\frac{1}{2}\beta\epsilon}}{2 \cosh(\frac{1}{2}\beta\epsilon)} = \frac{1}{1 + e^{-\beta\epsilon}} = \frac{1}{2} + \frac{1}{2} \tanh\left(\frac{1}{2}\beta\epsilon\right)$$

$$P(\text{excited}) = \bar{\rho}_{11} = \frac{e^{-\frac{1}{2}\beta\epsilon}}{2 \cosh(\frac{1}{2}\beta\epsilon)} = \frac{1}{1 + e^{\beta\epsilon}} = \frac{1}{2} - \frac{1}{2} \tanh\left(\frac{1}{2}\beta\epsilon\right)$$

<sup>6</sup>In machine learning, Boltzmann-distributed random variables are also known as *log-linear models*.

In terms of Pauli coordinates, the canonical density matrix is:

$$\bar{\rho} = \frac{1}{2} \begin{bmatrix} 1 + \tanh(\frac{1}{2}\beta\epsilon) & 0 \\ 0 & 1 - \tanh(\frac{1}{2}\beta\epsilon) \end{bmatrix} \Leftrightarrow \bar{\mathbf{r}} = \begin{bmatrix} 0 \\ 0 \\ \tanh(\frac{1}{2}\beta\epsilon) \end{bmatrix}$$

Thermal-equilibrium mean states are found on the  $z$  axis of the Bloch ball. The North pole is the “coldest” state: it has zero von Neumann entropy and is the limit of  $\bar{\mathbf{r}}$  as  $T \rightarrow 0^+$ . In the “hot” limit  $T \rightarrow \infty$ ,  $\bar{\mathbf{r}}$  approaches the maximum-entropy mean state  $\bar{\mathbf{r}} = \mathbf{0}$ . Figure 2.3 shows a plot of the von Neumann entropy of  $\bar{\rho}$  as a function of  $kT/\epsilon$ .

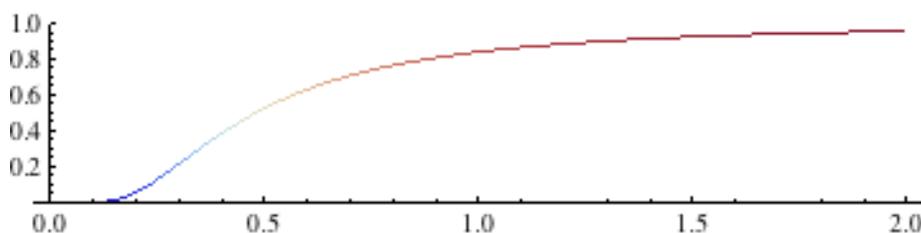


Figure 2.3: vN entropy (in bits) of the canonical density matrix as a function of  $kT/\epsilon$ .

### 2.5.1 Negative temperatures

Canonical density matrices can also be defined for negative temperatures. This counter-intuitive property can be explained more easily in terms of inverse temperature. A classical **thermodynamic macrostate** is an ordered triple  $(U, V, N)$ , where the letters represent internal energy, volume, and particle number.<sup>7</sup> Thermodynamic properties are determined by an entropy function  $S(U, V, N)$ . The classical definition of **temperature**  $T$  is:

$$kT \equiv \left( \frac{\partial S(U, V, N)}{\partial U} \right)^{-1}$$

Instead of temperature, define the **greed**  $\beta$  of a thermodynamic macrostate:

$$\beta \equiv \frac{\partial S(U, V, N)}{\partial U} = \frac{1}{kT}$$

Greed has units of bits per Joule (or nats per Joule, or digits per Joule). The name is a mnemonic based on an analogy from Schroeder’s *Thermal Physics* textbook.[22] A system with large positive  $\beta$  is “greedy for energy” and tends to acquire energy from its neighbors until all interacting systems have approximately equal greed. A system with negative  $\beta$  is “generous” and tends to give energy to its neighbors. (Schroeder jokingly refers to

<sup>7</sup>For systems with multiple particle types,  $N$  can be promoted to a vector.

negative- $\beta$  macrostates as *enlightened*.) Greed also replaces “absolute zero temperature” with “infinite greed,” which may be easier to interpret as a theoretical idealization.

The tanh function is odd:  $\tanh(-\frac{1}{2}\beta\epsilon) = -\tanh(\frac{1}{2}\beta\epsilon)$ . Every thermal qubit mixed state with  $\beta > 0$  thus has a generous counterpart found by reversing the sign of its  $\bar{z}$  coordinate. As  $\beta \rightarrow \infty$ ,  $\bar{\mathbf{r}}$  approaches the infinitely-greedy ground state, which is physically incapable of giving energy. Its infinitely-generous counterpart with  $\beta \rightarrow -\infty$  is the excited state, which cannot accept energy. The midpoint of these extremes is the maximum-entropy mixed state  $\bar{\mathbf{r}} = \mathbf{0}$ , which has  $\beta = 0$  and is neither greedy nor generous.

Figure 2.4 shows two plots of  $P(\text{excited})$  as a function of  $\beta\epsilon$  or  $kT/\epsilon$ . For negative temperatures, the probability of detecting the higher-energy result is *greater* than the maximum-entropy value of 50%. In this sense,  $T < 0$  should be interpreted as “hotter than  $T = \infty$ ,” which is counter-intuitive. An easier mnemonic is: thermal mixed states with  $\beta < 0$  have a surplus of energy, so they tend to give it away generously.

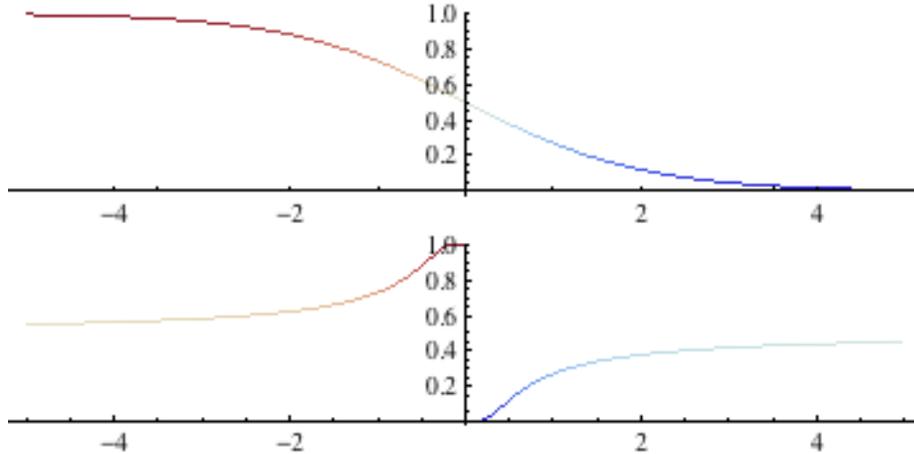


Figure 2.4: Probability of detecting excited energy  $+\frac{1}{2}\epsilon$  for a thermal-equilibrium qubit. Top:  $P(\text{excited})$  as function of  $\beta\epsilon$ . Bottom:  $P(\text{excited})$  as function of  $kT/\epsilon$ .

For an  $N$ -level quantum system, the maximum-entropy canonical density matrix is  $\bar{\rho} = \frac{1}{N}\hat{1}$ . The expectation value of an energy measurement is  $\langle H \rangle = \text{Tr}[\bar{\rho}\hat{H}] = \frac{1}{N}\text{Tr}[\hat{H}]$ . Any mixed state with  $\langle H \rangle > \frac{1}{N}\text{Tr}[\hat{H}]$  must *lose* energy to reach the maximum-entropy mixture. In Schroeder’s analogy, these systems tend to behave generously.

## 2.5.2 Binary entropy, logit, and logistic functions

Thermal qubit calculations can be simplified by using special functions related to the *logistic distribution*. From Section 2.4, the von Neumann entropy of a qubit mean state is:

$$S(\bar{\mathbf{r}}) = -\frac{1 + |\bar{\mathbf{r}}|}{2} \log\left(\frac{1 + |\bar{\mathbf{r}}|}{2}\right) - \frac{1 - |\bar{\mathbf{r}}|}{2} \log\left(\frac{1 - |\bar{\mathbf{r}}|}{2}\right)$$

Define the **binary entropy function**  $C(p)$ :<sup>8</sup>

$$C(p) \equiv -p \log(p) - (1-p) \log(1-p) \quad \text{for } 0 < p < 1$$

$C(p)$  is “coin-toss entropy,” i.e. the Shannon entropy of a random variable with two distinct outcomes, one of which occurs with probability  $p$ . To ensure compatibility with the Shannon-entropy convention  $0 \log(0) \rightarrow 0$ , define  $C(0)$  and  $C(1)$  to be zero. With this definition, the von Neumann entropy of a qubit mean state is:

$$S(\bar{\mathbf{r}}) = C\left(\frac{1}{2} + \frac{1}{2}|\bar{\mathbf{r}}|\right)$$

The derivative of the binary entropy function is (using base- $e$  for logarithms):

$$C'(p) = -\log(p) - 1 + \frac{1}{1-p} + \log(1-p) - \frac{p}{1-p} = \log\left(\frac{1-p}{p}\right)$$

This is the negative of the **logit function**, also called the **log-odds** of  $p$ :

$$\text{logit}(p) \equiv \log\left(\frac{p}{1-p}\right)$$

Figure 2.5 shows plots of  $C(p)$  and  $\text{logit}(p)$ .

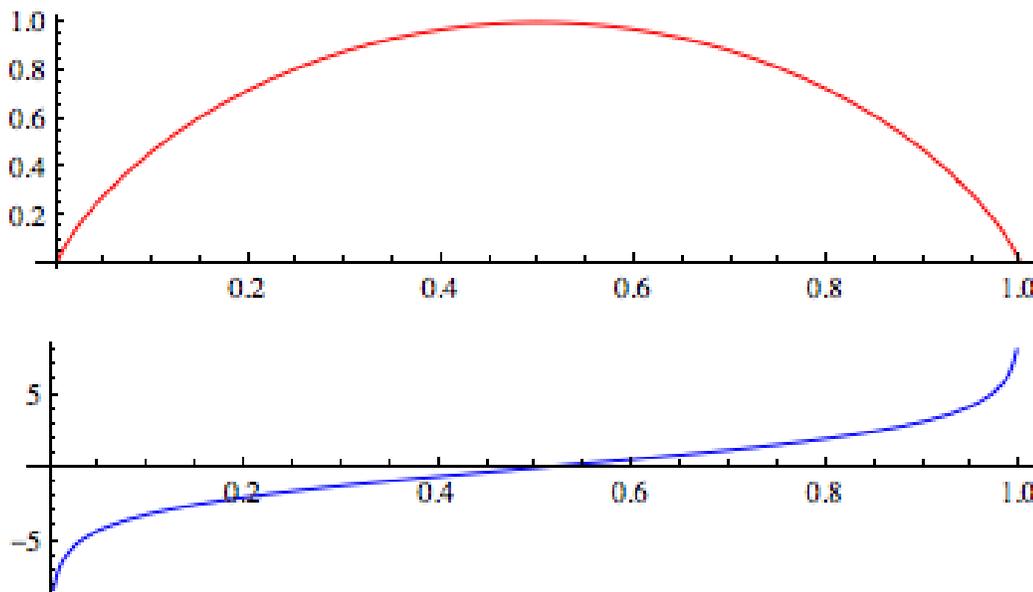


Figure 2.5: Binary entropy and logit functions in base-2.  
Top: binary entropy function  $C(p)$ . Bottom: logit function  $\text{logit}(p)$ .

<sup>8</sup>The notation  $H_2(p)$  is more common, but this conflicts with *Hamiltonian* and *heads*.

The relationship  $C'(p) = -\text{logit}(p)$  remains valid for other bases if  $C(p)$  and  $\text{logit}(p)$  are both defined using the same logarithm base:

$$C_b(p) \equiv -p \log_b(p) - (1-p) \log_b(1-p) \quad \text{logit}_b(p) \equiv \log_b\left(\frac{p}{1-p}\right)$$

If  $\text{logit}$  is defined using base  $e$ , then its inverse function is the **logistic function**:

$$\text{logistic}(x) \equiv \frac{e^x}{1+e^x} = \frac{1}{1+e^{-x}} = \frac{1}{2} + \frac{1}{2} \tanh\left(\frac{1}{2}x\right)$$

The logistic function is the cumulative distribution of the standard *logistic distribution*, and it is also the solution to the nonlinear *logistic differential equation* with initial value  $1/2$ . For the present purposes, it is useful because it provides a concise way to write thermal-equilibrium qubit mixed states. Using base- $e$ , the canonical qubit density matrix is:

$$\bar{\rho} = \frac{1}{2}\hat{1} + \frac{1}{2} \tanh\left(\frac{1}{2}\beta\epsilon\right)\hat{\sigma}_z = \begin{bmatrix} \text{logistic}(\beta\epsilon) & 0 \\ 0 & \text{logistic}(-\beta\epsilon) \end{bmatrix} = \text{logistic}(\beta\epsilon\hat{\sigma}_z)$$

Logistic and  $\text{logit}$  are inverse functions, so this result can be used to define the greed (or temperature) of any diagonal qubit density matrix as a function of its  $z$ -coordinate:

$$\beta\epsilon = \text{logit}(\bar{\rho}_{00}) = \text{logit}\left(\frac{1}{2} + \frac{1}{2}z\right)$$

### 2.5.3 MaxEnt mean states

The canonical qubit density matrix can be derived using a version of the **maximum entropy principle**. This method calculates the mean state which maximizes von Neumann entropy subject to a constraint on the expected energy of the qubit.

As before, consider a qubit with reference Hamiltonian  $\hat{H} = -\frac{1}{2}\epsilon\hat{\sigma}_z$ . The energy expectation value of a state with Pauli coordinates  $(x, y, z)$  is  $\langle H \rangle = \text{Tr}[\hat{\rho}\hat{H}] = -\frac{1}{2}\epsilon z$ . For any  $U \in [-\frac{1}{2}\epsilon, \frac{1}{2}\epsilon]$ , define a **macrostate**  $\Omega_U$  as the set of all pure qubit states whose energy expectation value equals  $U$ . For notational clarity, define a dimensionless energy  $u \equiv U/\epsilon$ . The macrostate  $\Omega_U$  is then the set of all Bloch vectors  $\mathbf{r}$  such that

$$z = -2u \quad \text{and} \quad |\mathbf{r}|^2 = x^2 + y^2 + 4u^2 = 1$$

Unless  $U = \pm\frac{1}{2}\epsilon$ , many pure states are consistent with the constraint  $\langle H \rangle = U$ . Represent the true state  $\mathbf{r}$  as a random variable with sample space  $\Omega_U$ . What probability distribution should be assigned to  $\mathbf{r}$ ? The MaxEnt principle suggests choosing a distribution which

maximizes  $S(\bar{\mathbf{r}})$ .  $S(\bar{\mathbf{r}})$  is a function of mean-state radius  $|\bar{\mathbf{r}}|$  only, and it is maximized by minimizing  $|\bar{\mathbf{r}}|$ . The maximum-entropy mean state with  $\bar{z} = -2u$  is therefore:

$$\bar{\mathbf{r}} = (0, 0, -2u)$$

Many distributions are consistent with this mean state. For example, one could assign a distribution with two equally-probable outcomes  $\mathbf{r} = (\pm\sqrt{1-4u^2}, 0, -2u)$ , or a continuous distribution which is uniform in  $\phi$  with fixed  $\theta = \arccos(-2u)$ . But regardless of the distribution used to calculate them, MaxEnt mean states are located on the  $z$ -axis.

Define an entropy function  $S$  which inputs  $U$  and returns the von Neumann entropy of the MaxEnt mean state with  $\langle H \rangle = U$ . In terms of the binary entropy function  $C$ ,

$$S(U) = C\left(\frac{1}{2} + \frac{1}{2}|\bar{\mathbf{r}}|\right) = C\left(\frac{1}{2} + \frac{1}{2}|\bar{z}|\right) = C\left(\frac{1}{2} + \frac{1}{\epsilon}|U|\right)$$

Define the greed of a MaxEnt mean state as the  $U$  derivative of its entropy function:

$$\beta \equiv \frac{\partial S}{\partial U} = S'(U) = \begin{cases} \frac{1}{\epsilon} \text{logit}\left(\frac{1}{2} + \frac{1}{\epsilon}|U|\right) & \text{if } U < 0 \\ \frac{1}{\epsilon} \text{logit}\left(\frac{1}{2}\right) = 0 & \text{if } U = 0 \\ -\frac{1}{\epsilon} \text{logit}\left(\frac{1}{2} + \frac{1}{\epsilon}|U|\right) & \text{if } U > 0 \end{cases}$$

Using the symmetry property  $-\text{logit}\left(\frac{1}{2} + u\right) = \text{logit}\left(\frac{1}{2} - u\right)$  (see Figure 2.5), these results can be consolidated into one formula:

$$\beta\epsilon = \text{logit}\left(\frac{1}{2} - \frac{1}{\epsilon}U\right) = \text{logit}\left(\frac{1}{2} + \frac{1}{2}z\right) \Leftrightarrow \text{logistic}(\beta\epsilon) = \frac{1}{2} + \frac{1}{2}z$$

The unique MaxEnt mean state which satisfies the constraint  $S'(U) = \beta$  must be:

$$\bar{\rho} = \begin{bmatrix} \text{logistic}(\beta\epsilon) & 0 \\ 0 & \text{logistic}(-\beta\epsilon) \end{bmatrix}$$

which is the canonical qubit density matrix.

### 3 RANDOM QUBITS

The examples in this chapter are intended to introduce drunk models without the distracting technical difficulties of solving stochastic differential equations on manifolds. These models are too simple to realistically describe qubit decoherence, but they illustrate the physical assumptions of the stochastic models in Chapter 4. In particular, each demonstrates how unitary evolution can *appear* to turn pure states into mixed states if one does not carefully distinguish between true states, mean states, and estimated states.

The **quantum Loschmidt paradox** echoes Loschmidt’s criticism of Boltzmann’s H-theorem: how can reversible laws of physics lead to irreversible time evolution?<sup>1</sup> If time evolution of quantum systems is unitary, then unobserved states evolve reversibly and von Neumann entropy is constant in time. Does von Neumann entropy disobey the Second Law of Thermodynamics, or do states evolve in a non-unitary way?

For the thought experiments in this chapter, the answer is “none of the above.” In each example, an initial pure state appears to gain von Neumann entropy. The paradoxes are resolved by defining entropy exclusively for *random variables*, not for *physical systems*. This semantic convention is closely related to Jaynes’ and Wigner’s characterization of entropy as “an anthropomorphic concept,” which is discussed in Chapter 7.[24]

The *Allyson’s Choice* and *Decoherence by 1000 Small Cuts* thought experiments originally appeared in [25]. They were designed to show that, at least for some experiments, laboratory noise reversibly encrypts quantum information rather than irreversibly destroying it.

In the *Zech’s Qubit* thought experiment, a qubit state is altered by a noisy classical field. If the field is measured precisely by some external equipment, then it is possible to calculate the true state and return the qubit to its initial state using only unitary operations. In this case, the effects of decoherence can be “undone” in a way which is conceptually similar to more sophisticated *engineered decoherence* experiments such as [26].

---

<sup>1</sup>The quantum Loschmidt paradox should not be confused with the cosmological time-reversal paradox, which may or may not be related. See e.g. [23] for a discussion of both.

### 3.1 Allyson’s choice

Professor Bob intends to replicate a *welcher-weg* experiment to demonstrate de Broglie interference for his students. A particle launcher sends many identical neutrons, one at a time, through the Mach-Zender interferometer shown in Figure 3.1. Two physically-separated detectors are each designed to click if a neutron is detected. According to quantum theory, the detection probabilities depend sinusoidally on  $\varphi$ , the difference of phases acquired by the neutron’s wavefunction along either path.

Bob’s graduate student Allyson plans a prank designed to make Bob’s data appear consistent with classical mechanics. Before each neutron is launched, she flips a fair coin, records the result, and saves it to a file on a flash-memory stick. If the coin lands heads-up, then she runs the experimental trial as Bob intended. If tails, then she covertly reverses the orientation of the second beamsplitter.<sup>2</sup> If the ratio of heads/tails results is  $\approx 1$ , then Bob’s detector counts will appear to be independent of  $\varphi$ .

Bob erroneously concludes that the neutrons in the experiment have decohered into classical states. From Bob’s point of view, the neutrons’ evolution appears irreversible: the state of each neutron is pure before passing through the interferometer and mixed afterward. Contrariwise, Allyson represents each neutron using pure states only. She plans to reveal the prank to Bob by sorting the data into “heads only” and “tails only” sets, each of which clearly shows  $\varphi$ -dependent detector counts. From Allyson’s point of view, evidence of de Broglie interference was not destroyed – it was merely hidden from Bob.

Unfortunately, Allyson misplaces her memory stick before revealing the prank. Without knowing which trials were “heads” results and which were “tails,” she cannot unscramble Bob’s data. Though she cannot calculate the true state of the neutron on any given trial, she can still represent each coin toss as a random variable and calculate a mean state. The coin-toss history file then becomes a literal example of missing information.

Bob’s intended experiment is a simplified version of experiments such as [27].<sup>3</sup> The neutron’s true state is assumed to evolve by unitary transformations on each trial. Neutron interferometry is used because it is a well-researched topic which is relatively easy to visualize. Similar experiments with different interferometer configurations have been performed using sodium atoms, C<sub>60</sub> buckyballs, and even larger molecules.[30][31][32]

<sup>2</sup>Adjusting a beamsplitter thousands of times without one’s advisor noticing may be impractical. More realistically, imagine the second splitter’s orientation is controlled by a computer. Allyson secretly alters its control software to choose one of two equally-probable pseudo-random outcomes before each trial.

<sup>3</sup>The experiment also resembles Wheeler’s delayed-choice experiment as described in [28] and realized in [29]. Wheeler’s primary question – “Can information propagate faster than light?” – is avoided by replacing the photon with a nonrelativistic neutron with kinetic energy  $\ll m_n c^2$ .

Mathematically-identical experiments which use different qubit designs are also possible. For examples using superconducting qubits, see e.g. [33] and [34]. Subsection 3.2.2 rewrites the experiment in terms of Bloch-sphere rotations for a generic qubit.

### 3.1.1 Bob's intended experiment

On each trial, a single neutron is sent through a lossless beamsplitter  $S_1$ . Two mirrors  $M_L, M_R$ , a second splitter  $S_2$ , and two detectors  $D_L, D_R$  are placed as shown in Figure 3.1. At each stage, the two paths available to the neutron are arbitrarily named  $|L\rangle$  and  $|R\rangle$ . Each neutron can be detected by either  $D_L$  or  $D_R$ , but not both.

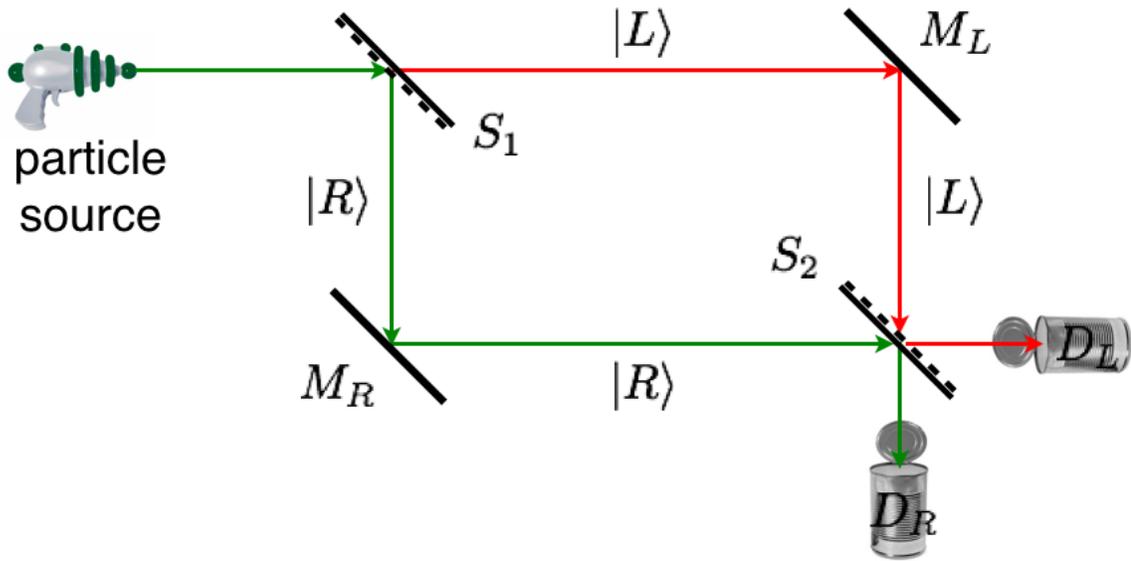


Figure 3.1: Mach-Zender interferometer for Bob's intended *welcher-weg* experiment.  $S_1, S_2$  are lossless beamsplitters.  $M_L, M_R$  are mirrors.  $D_L, D_R$  are detectors.

Classically, one expects the neutron to either reflect from or transmit through each splitter. For example, the neutron might reflect from  $S_1$ , then transmit through  $S_2$  to be detected by  $D_L$ . The two splitters produce four mutually-exclusive classical possibilities:

$$\text{start} \rightarrow |R\rangle \rightarrow D_R \quad \text{start} \rightarrow |R\rangle \rightarrow D_L \quad \text{start} \rightarrow |L\rangle \rightarrow D_R \quad \text{start} \rightarrow |L\rangle \rightarrow D_L$$

Quantum mechanics allows each splitter to send the neutron to a superposition of transmitted and reflected paths. Represent any such superposition with a column vector:

$$|\Psi\rangle = \alpha|L\rangle + \beta|R\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

Following the advice of Zeilinger in [35], assume the action of each splitter on the column vector  $|\Psi\rangle$  can be represented by a  $2\times 2$  special unitary matrix:

$$\hat{S} = \begin{bmatrix} r_{LL} & t_{LR} \\ t_{RL} & r_{RR} \end{bmatrix} \quad r_{RR} = r_{LL}^*, \quad t_{RL} = -t_{LR}^*, \quad \text{Det}[\hat{S}] = r_{RR}r_{LL} - t_{RL}t_{LR} = 1$$

The numbers  $r_{LL}, r_{RR}, t_{LR}, t_{RL}$  are (complex) reflection and transmission coefficients. Bob chooses to use splitters  $\hat{S}_1$  and  $\hat{S}_2$  as follows:

$$\hat{S}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad \hat{S}_2 = (\hat{S}_1)^{-1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

The first splitter has equal transmission and reflection probabilities, and it produces no phase shift along either path. Let  $|\Psi_0\rangle = |R\rangle$  be the initial state of the neutron. Its state after passing through the first splitter  $\hat{S}_1$  is:

$$\hat{S}_1|\Psi\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Bob orients the second splitter such that  $\hat{S}_2 = \hat{S}_1^{-1} = \hat{S}_1^\dagger$ .

As the neutron travels, its wavefunction acquires a phase shift.<sup>4</sup> The overall phase of  $|\Psi\rangle$  is unobservable, but the phase difference between different paths can affect detection probabilities. Represent the phase shifts with a diagonal unitary operator  $\hat{\Phi}$ . The composite operation “do  $\hat{S}_1$ , then  $\hat{\Phi}$ , then  $\hat{S}_2$ ” is represented by the operator product  $\hat{H} \equiv \hat{S}_2\hat{\Phi}\hat{S}_1$ . (The notation  $\hat{H}$  is chosen to suggest *heads*, not *Hamiltonian*.)

$$\hat{\Phi} \equiv \begin{bmatrix} e^{i\theta_L} & 0 \\ 0 & e^{i\theta_R} \end{bmatrix} \quad \hat{H} = \hat{S}_2\hat{\Phi}\hat{S}_1 = e^{i\theta_L} \frac{1}{2} \begin{bmatrix} 1 + e^{i\varphi} & 1 - e^{i\varphi} \\ 1 - e^{i\varphi} & 1 + e^{i\varphi} \end{bmatrix} \quad \varphi \equiv \theta_R - \theta_L$$

If the neutron’s initial state is  $|R\rangle$ , then it reaches the detectors in state  $\hat{H}|R\rangle$ :

$$|\Psi_H\rangle \equiv \hat{H}|R\rangle = e^{i\theta_L} \frac{1}{2} \begin{bmatrix} 1 + e^{i\varphi} & 1 - e^{i\varphi} \\ 1 - e^{i\varphi} & 1 + e^{i\varphi} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = e^{i\theta_L} \frac{1}{2} \begin{bmatrix} 1 - e^{i\varphi} \\ 1 + e^{i\varphi} \end{bmatrix}$$

<sup>4</sup>Depending on the details of the experiment, the physical source of this phase shift might be e.g. a gravitational or magnetic potential [27] or pure de Broglie interference.[32]

The probabilities of the neutron being detected by  $D_L$  and  $D_R$  are:

$$P(D_L) = |\langle L|\hat{H}|R\rangle|^2 = \left|\frac{1}{2}(1 - e^{i\varphi})\right|^2 = \frac{1}{2}(1 - \cos \varphi)$$

$$P(D_R) = |\langle R|\hat{H}|R\rangle|^2 = \left|\frac{1}{2}(1 + e^{i\varphi})\right|^2 = \frac{1}{2}(1 + \cos \varphi)$$

Bob's intended experimental procedure is:

1. For each trial, send one neutron through the interferometer.
2. Every 1000 trials, adjust the path lengths such that  $\varphi$  advances by some amount  $\Delta$ .
3. Plot the number of  $D_L$  and  $D_R$  counts for each value of  $\varphi$ .

Fig. 3.2 shows simulated detector counts with  $\Delta = \frac{\pi}{8}$ .

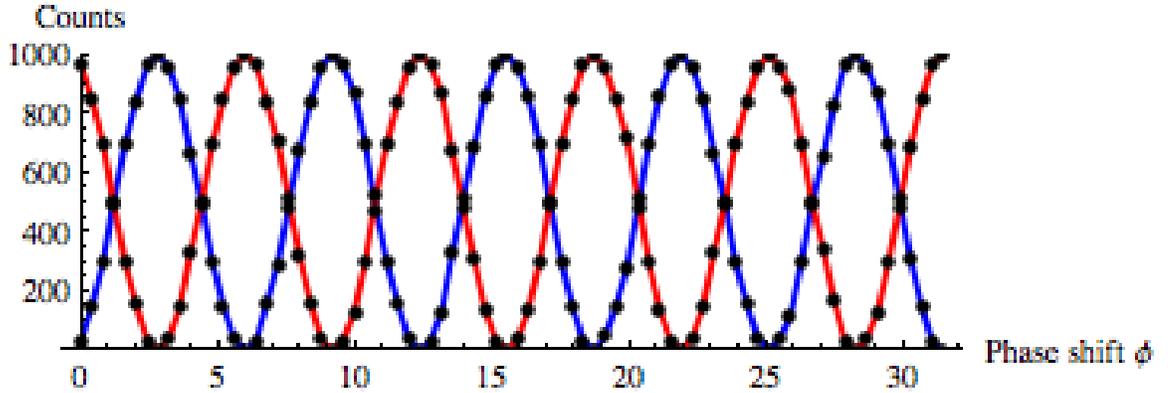


Figure 3.2: Bob's intended histogram.

Horizontal axis is phase shift  $\varphi$  between paths. Blue:  $D_L$  counts. Red:  $D_R$  counts.

### 3.1.2 Allyson's randomized experiment

On any trial for which Allyson's coin lands heads, she reverses the orientation of  $\hat{S}_2$  so that its matrix representation is  $\hat{S}_2^T = \hat{S}_2^\dagger = \hat{S}_1$ . The action of the M-Z apparatus on  $|\Psi\rangle$  is then represented by a unitary "tails" operator  $\hat{T} = \hat{S}_1 \hat{\Phi} \hat{S}_1$ :

$$\hat{T} \equiv \hat{S}_1 \hat{\Phi} \hat{S}_1 = e^{i\theta_L} \frac{1}{2} \begin{bmatrix} 1 - e^{i\varphi} & 1 + e^{i\varphi} \\ -1 - e^{i\varphi} & -1 + e^{i\varphi} \end{bmatrix}$$

For tails trials, she calculates that the state vector immediately prior to detection is  $\hat{T}|R\rangle$ :

$$|\Psi_T\rangle \equiv \hat{T}|R\rangle = e^{i\theta_L} \frac{1}{2} \begin{bmatrix} 1 + e^{i\varphi} \\ -1 + e^{i\varphi} \end{bmatrix}$$

If the coin lands tails, then she calculates the detector probabilities to be:

$$P(D_L|\text{tails}) = \|\langle L|\hat{T}|R\rangle\|^2 = \|\frac{1}{2}(1 + e^{i\varphi})\|^2 = \frac{1}{2}(1 + \cos \varphi)$$

$$P(D_R|\text{tails}) = \|\langle R|\hat{T}|R\rangle\|^2 = \|\frac{1}{2}(-1 + e^{i\varphi})\|^2 = \frac{1}{2}(1 - \cos \varphi)$$

If the coin lands heads, then she predicts the same detection probabilities as Bob:

$$P(D_L|\text{heads}) = \|\langle L|\hat{H}|R\rangle\|^2 = \|\frac{1}{2}(1 - e^{i\varphi})\|^2 = \frac{1}{2}(1 - \cos \varphi)$$

$$P(D_R|\text{heads}) = \|\langle R|\hat{H}|R\rangle\|^2 = \|\frac{1}{2}(1 + e^{i\varphi})\|^2 = \frac{1}{2}(1 + \cos \varphi)$$

The total probabilities for each detector are found by the Law of Total Probability:

$$P(D_L) = P(D_L|\text{heads})P(\text{heads}) + P(D_L|\text{tails})P(\text{tails}) = \frac{1}{2}$$

$$P(D_R) = P(D_R|\text{heads})P(\text{heads}) + P(D_R|\text{tails})P(\text{tails}) = \frac{1}{2}$$

Figure 3.3 shows results from a simulation of Allyson's randomized experiment.

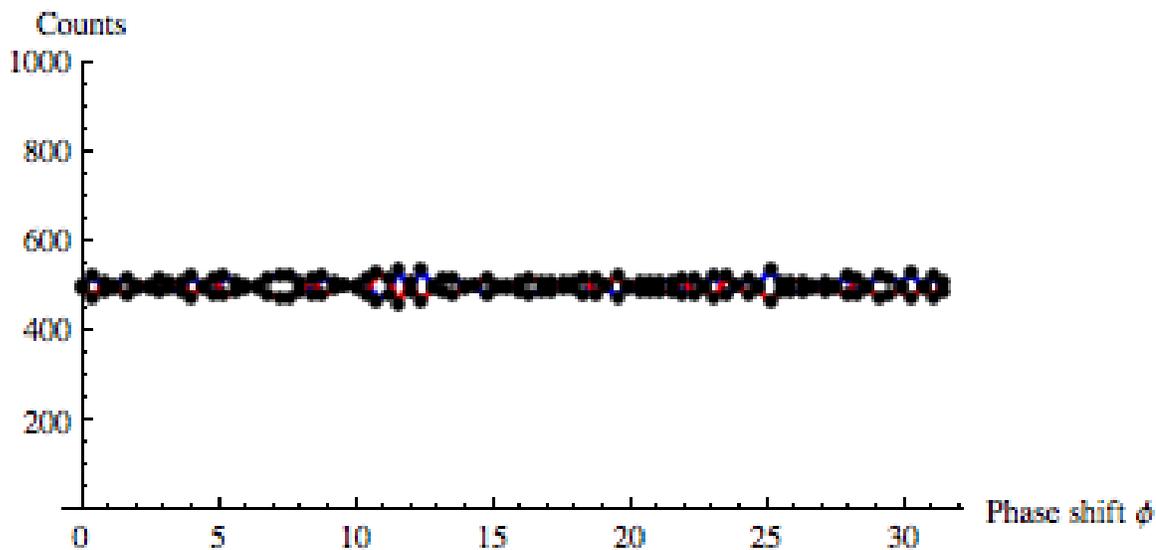


Figure 3.3: Bob's sabotaged results.

Horizontal axis is phase shift  $\varphi$  between paths. Blue:  $D_L$  counts. Red:  $D_R$  counts.

Allyson's shenanigans have concealed all evidence of de Broglie interference. From Bob's point of view, each neutron appears to have decohered to a classical state during its journey through the interferometer. Bob then faces a quantum Loschmidt paradox: the neutron evolved only by passing through unitary beamsplitters and reflecting off perfect mirrors, so its evolution must have been reversible. Where did the information go?

In this case, the resolution is simple: Bob's assumption of identical trials is simply *wrong*. Allyson knows that the neutron evolved reversibly and unitarily during each trial. Using

her coin-toss history file, she writes a program to unscramble the data. The program finds all trials marked “tails,” stores them separately from the “heads” trials, and produces a histogram for each set of trials. Pseudocode is shown below:

```

for each phi:
  for each trial:
    if coin was heads:
      number_of_heads += 1
      if neutron was detected on left:
        heads_wins[phi] += 1
    else if coin was tails:
      number_of_tails += 1
      if neutron was detected on left:
        tails_wins[phi] += 1
plot histogram of (heads_wins) and (number_of_heads - heads_wins)
plot histogram of (tails_wins) and (number_of_tails - tails_wins)

```

Examples of this program’s output are shown in Figure 3.4.

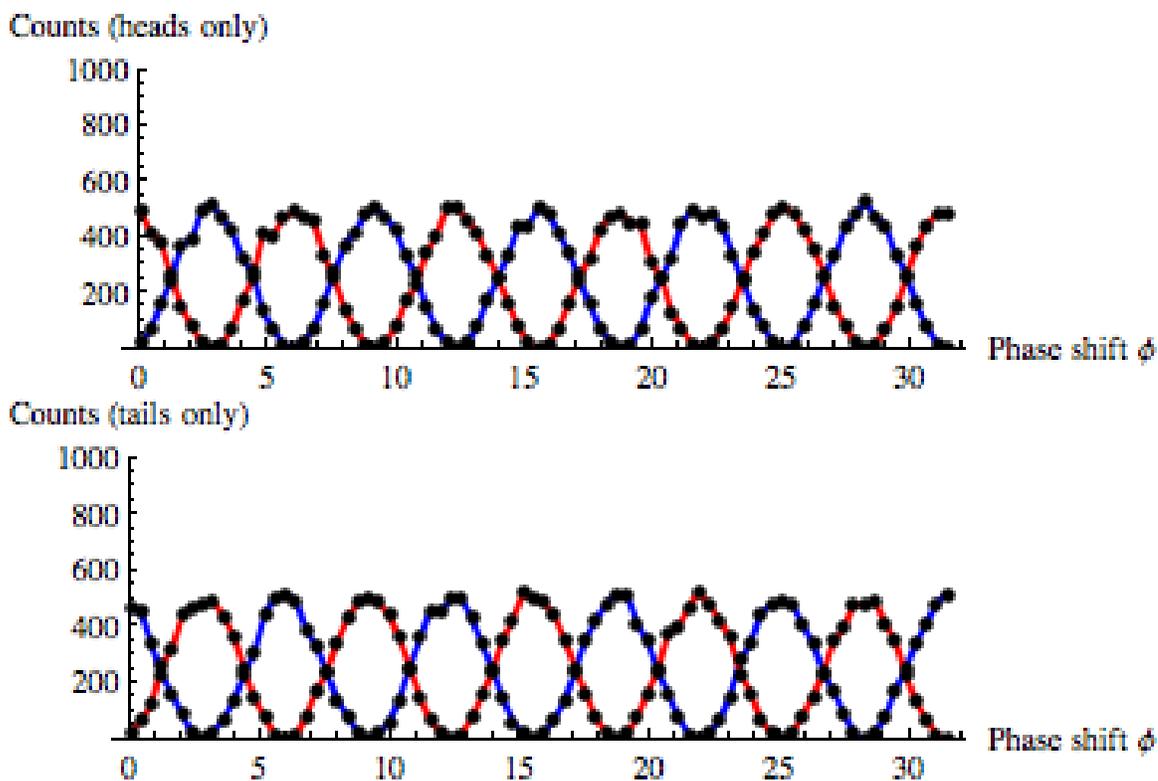


Figure 3.4: Allyson’s decrypted results. Top: heads trials only. Bottom: tails trials only. Horizontal axis is phase shift  $\varphi$  between paths. Blue:  $D_L$  counts. Red:  $D_R$  counts.

### 3.1.3 Information was encrypted, not destroyed

Even if Bob discovers Allyson's subterfuge, he cannot unscramble the data without knowing the results of the coin flips. The data has effectively been encrypted with a provably-secure *one-time pad*, and the coin-toss history is its password.<sup>5</sup> As an example, suppose the first 8 coin toss results were *THHT THTH*. The first 8 states must have been:

$$|\Psi_T\rangle, |\Psi_H\rangle, |\Psi_H\rangle, |\Psi_T\rangle, |\Psi_T\rangle, |\Psi_H\rangle, |\Psi_T\rangle, |\Psi_H\rangle$$

Represent Allyson's list of states, Bob's list of intended states, and the coin-toss results each as a binary string with 0 for heads and 1 for tails. Bob's list is then a plaintext composed entirely of zeros, Allyson's list is a ciphertext, and the coin-toss history is a password. Table 3.1 shows the encryption scheme, which is a **binary Vernam cipher**: add (mod 2) each element in the plaintext to the corresponding element in the password. The password meets all the requirements of a one-time pad: it is the same length as the plaintext, used only once, and each bit is chosen randomly with 0 and 1 equally probable.

Table 3.1: Example cipher for coin history *THHT THTH*.

Bob's plaintext	0	0	0	0	0	0	0	0
+ Coin result (mod 2)	1	0	0	1	1	0	1	0
= Allyson's ciphertext	1	0	0	1	1	0	1	0

Now suppose that before decrypting the data, Allyson misplaces her coin-toss history file. Without the password, neither Allyson nor Bob knows what state was actually present on any given trial. But not *all* information about the experiment is lost: Allyson knows each state was either  $|\Psi_H\rangle$  or  $|\Psi_L\rangle$ . For each of these states, Allyson constructs a pure density matrix  $|\Psi\rangle\langle\Psi|$ . She can now represent the true state as a density-matrix-valued random variable with two equally-probable outcomes  $\hat{\rho}_H$  and  $\hat{\rho}_T$ :

$$\hat{\rho}_H \equiv |\Psi_H\rangle\langle\Psi_H| = \frac{1}{4} \begin{bmatrix} 1 - e^{i\varphi} \\ 1 + e^{i\varphi} \end{bmatrix} \begin{bmatrix} 1 - e^{-i\varphi} & 1 + e^{-i\varphi} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 - \cos(\varphi) & -i \sin(\varphi) \\ i \sin(\varphi) & 1 + \cos(\varphi) \end{bmatrix}$$

$$\hat{\rho}_T \equiv |\Psi_T\rangle\langle\Psi_T| = \frac{1}{4} \begin{bmatrix} 1 + e^{i\varphi} \\ -1 + e^{i\varphi} \end{bmatrix} \begin{bmatrix} 1 + e^{-i\varphi} & -1 + e^{-i\varphi} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 + \cos(\varphi) & -i \sin(\varphi) \\ i \sin(\varphi) & 1 - \cos(\varphi) \end{bmatrix}$$

<sup>5</sup>See [36] for the relevant definitions and Shannon's security proof. Kotelnikov independently proved a similar result in a classified report.[37]

The mean state is the expectation value of the true state:

$$\bar{\rho} \equiv E[\hat{\rho}] = \frac{1}{2}\hat{\rho}_H + \frac{1}{2}\hat{\rho}_T = \frac{1}{2} \begin{bmatrix} 1 & -i \sin(\varphi) \\ i \sin(\varphi) & 1 \end{bmatrix}$$

The diagonal elements of  $\bar{\rho}$  are the total probabilities of  $D_L$  or  $D_R$  detecting the neutron. The mean state predicts both probabilities are  $\frac{1}{2}$  regardless of the phase difference  $\varphi$ . If Bob's skeptical students argue that quantum theory is wrong and the neutrons behaved classically, Allyson and Bob will be unable to prove otherwise without finding the missing coin-toss history or performing another experiment. By misplacing the password, Allyson has effectively encrypted the data and thrown away the key.

Eigenvalues of  $\bar{\rho}$  are  $\lambda_{\pm} = \frac{1}{2}[1 \pm \sin(\varphi)]$ . The mean state is pure only if the eigenvalues are 1 and 0, which occurs only on trials for which  $|\sin(\varphi)| = 1$ . Figure 3.5 shows the von Neumann entropy of the mean state, which is given by the following unwieldy expression:

$$S[\bar{\rho}] = -[\lambda_+ \log(\lambda_+) + \lambda_- \log(\lambda_-)] = \frac{1}{2} \left[ \log \left( \frac{4}{\cos^2[\varphi]} \right) + \log \left( \frac{1 - \sin[\varphi]}{1 + \sin[\varphi]} \right) \sin[\varphi] \right]$$

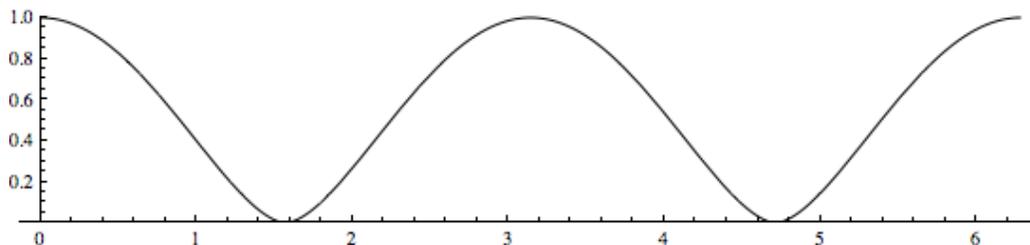


Figure 3.5: Von Neumann entropy (bits) of mean state as a function of  $\varphi$ .

Each mean state can be given a Fano-style interpretation as a “state of less than maximum information.” For each trial, Allyson accidentally discarded 1 bit of information: the answer to the question “heads or tails?” For trials in which  $|\sin(\varphi)| \approx 1$ , this information is mostly irrelevant;  $\hat{\rho}_H \approx \hat{\rho}_T$ , so the neutron’s state can be calculated approximately without knowing the result of the coin toss for that trial. When  $\sin(\varphi) \approx 0$ , the mean state is near the maximum-entropy mixture  $\frac{1}{2}\hat{1}$ , and its von Neumann entropy is  $S[\bar{\rho}] \approx 1$  bit. In these cases, neither Allyson nor Bob has any idea what state was actually present.

In this thought experiment, irreversibility is an accidental inconvenience, not a fundamental law of physics. By assumption, each neutron’s state vector evolved unitarily and reversibly on each trial; it is only the mean state  $\bar{\rho}$  which has gained entropy. Even this evolution is not absolutely irreversible: if Allyson finds the missing memory stick, then she can decrypt the data and avoid the need to use a mixed-state representation. Information is “missing” from Allyson’s and Bob’s description of the experiment, not from the universe.

## 3.2 Decoherence by 1000 small cuts

In a real laboratory, Nature often plays the role of saboteur, and Nature generally does not bother saving a detailed history of its actions to a file on a memory stick. This thought experiment assumes Allyson and Bob collaborate perfectly, but their laboratory does not. Bob repeats his neutron-interferometry experiment, and this time Allyson cooperates fully. This time, their control of the interferometer is imperfect in a very specific way: the phase difference  $\varphi$  between the neutron paths varies erratically by an amount that is not negligible but is impractical to measure directly.

The randomness in this thought experiment is contrived for mathematical convenience. The physical source of imprecision in  $\varphi$  is left to readers' imaginations; perhaps it is seismic vibrations, thermal expansion of the hardware, or some other nuisance. For a realistic description of phase noise in Mach-Zehnder matter interferometers, see e.g. [38].

Allyson and Bob calculate that the state immediately before detection will be:

$$\hat{\rho} = |\Psi_H\rangle\langle\Psi_H| = \frac{1}{2} \begin{bmatrix} 1 - \cos(\varphi) & -i \sin(\varphi) \\ i \sin(\varphi) & 1 + \cos(\varphi) \end{bmatrix}$$

To reduce fluctuations, they decide to perform 1000 trials for each value of  $\varphi$ . In an ideal experiment, the Law of Large Numbers implies that the ratio  $D_L/D_R$  of detection counts will, with high probability, be close to its expectation value  $|\langle L|\Psi\rangle|^2 / |\langle R|\Psi\rangle|^2$ . Unfortunately, imprecise control of  $\varphi$  means  $\hat{\rho}$  is not identical for all trials.

### 3.2.1 A convoluted model

To account for errors in  $\varphi$ , Allyson and Bob represent  $\varphi$  as a random variable which is identically *distributed* over all 1000 trials. Neither scientist can calculate the precise true state of the neutron on any trial, but they can assign  $\varphi$  a probability distribution. As an approximation, they assign  $\varphi$  a normal distribution with mean  $\mu$  and variance  $\sigma^2$ . For calculational simplicity, suppose they have near-perfect control of  $\mu$ , but  $\sigma^2$  is some fixed positive number. For each value of  $\mu$ , they calculate a mean state  $\bar{\rho}$ :

$$\bar{\rho} \equiv E[\hat{\rho}] = \int_{-\infty}^{\infty} \hat{\rho}(\varphi) p(\varphi) d\varphi = \int_{-\infty}^{\infty} \frac{1}{2} \begin{bmatrix} 1 - \cos(\varphi) & -i \sin(\varphi) \\ i \sin(\varphi) & 1 + \cos(\varphi) \end{bmatrix} \frac{e^{-\frac{1}{2}\left(\frac{\varphi-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}} d\varphi$$

Each matrix element is a convolution.<sup>6</sup> The resulting mean state is:

$$\bar{\rho} = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} -\cos(\mu) & -i \sin(\mu) \\ i \sin(\mu) & \cos(\mu) \end{bmatrix} e^{-\frac{1}{2}\sigma^2}$$

This mean state predicts the following detection probabilities:

$$P(D_L) = \frac{1}{2} - \frac{1}{2} \cos(\mu) e^{-\frac{1}{2}\sigma^2} \quad P(D_R) = \frac{1}{2} + \frac{1}{2} \cos(\mu) e^{-\frac{1}{2}\sigma^2}$$

If Bob wants to convince his students that de Broglie interference is a real phenomenon, he and Allyson must keep the variance of  $\varphi$  as small as possible. In the limit  $\sigma^2 \rightarrow 0$ ,  $\bar{\rho}$  is pure. If  $0 < \sigma^2 \ll 1$ , then Bob's plot is similar to what he expected: sinusoidal oscillations with slightly reduced amplitude. But as  $\sigma^2$  increases, the amplitude of oscillations decreases exponentially. If  $\sigma^2 \gg 1$ , then  $\bar{\rho} \approx \frac{1}{2} \hat{1}$  and the predicted sinusoidal phase dependence is hopelessly scrambled by the cumulative effect of a thousand small experimental errors.

Quantum computing experimenters will probably not be surprised by the conclusion that noisy experiments can produce classical-looking results. But it may be surprising that any credible prediction at all can be made from such a crudely oversimplified model. The model used here makes no attempt to describe the laboratory environment or any extra degrees of freedom for the neutron – it only assumes that on each trial,  $\varphi$  is a random variable which is i.i.d. normal with mean  $\mu$  controlled by experimenters and constant variance  $\sigma^2$ .

The point is: not-quite-identical trials can corrupt a quantum experiment such that physical states *appear* to decohere. Uncontrolled interactions “hide” quantum information by effectively encrypting the sequence of final states. If a complete history of all interactions in an experiment is not recorded, then it may be impossible to decrypt the resulting data. A mixed-state representation of the system may become a practical necessity even if the true state vector evolved by unitary transformations on every trial.

### 3.2.2 Generalization to other qubit designs

The *Allyson's Choice* and *1000 Cuts* experiments use neutron interferometry – but in principle, both experiments can be realized using any qubit design. (In practice, doing so may require substantial experimental skill.) One simply redefines  $|L\rangle$  and  $|R\rangle$  to be the ground and excited energy eigenstates  $|0\rangle$  and  $|1\rangle$ . The next step is to find some physical processes which alter the qubit state  $|\Psi\rangle$  in the same way the unitary operators  $\hat{S}_1$  and  $\hat{\Phi}$  do. The

<sup>6</sup>Here  $\varphi \in (-\infty, \infty)$ . One could instead restrict  $\varphi$  to the interval  $[0, 2\pi)$ , in which case the probability density  $p(\varphi)$  becomes a *wrapped Gaussian* and the integrals are *circular convolutions*.

result of acting  $\hat{S}_1$  on a generic qubit (pure or mixed) state  $\hat{\rho}$  is  $\hat{S}_1|\Psi\rangle\langle\Psi|\hat{S}_1^\dagger$ :

$$|\Psi\rangle \rightarrow \hat{S}_1|\Psi\rangle \quad |\Psi\rangle\langle\Psi| \rightarrow \hat{S}_1|\Psi\rangle\langle\Psi|\hat{S}_1^\dagger \quad \hat{\rho} \rightarrow \hat{S}_1\hat{\rho}\hat{S}_1^\dagger$$

Written as a real linear combination of  $\hat{1}$  and Pauli matrices, the transformed state is:

$$\begin{aligned} \hat{S}_1\hat{\rho}\hat{S}_1^\dagger &= \frac{1}{4} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1+z & x-iy \\ x+iy & 1-z \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1+x & -z-iy \\ -z+iy & 1-x \end{bmatrix} \\ &= \frac{1}{2} \left( \hat{1} - z\hat{\sigma}_x + y\hat{\sigma}_y + x\hat{\sigma}_z \right) \end{aligned}$$

An initial state  $\mathbf{r} = (x, y, z)$  is transformed to  $(-z, y, x)$ . Let  $\hat{S}_1$  denote the matrix representing  $\hat{S}_1$  in Pauli coordinates. This matrix must be:

$$\hat{S}_1 = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \hat{S}_1\mathbf{r} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -z \\ y \\ x \end{bmatrix}$$

$\hat{S}_1$  can be visualized as a  $-90^\circ$   $y$ -axis rotation on the Bloch sphere. (The sign convention used here is: a positive  $x$  rotation moves  $\mathbf{r}$  clockwise as viewed by someone looking in the  $+x$  direction. Similar conventions are used for  $y$  and  $z$ .)

The action of  $\hat{\Phi}$  on a generic qubit pure state is  $\hat{\rho} \rightarrow \hat{\Phi}\hat{\rho}\hat{\Phi}^\dagger$ :

$$\hat{\Phi}\hat{\rho}\hat{\Phi}^\dagger = \frac{1}{2} \begin{bmatrix} e^{i\theta_L} & 0 \\ 0 & e^{i\theta_R} \end{bmatrix} \begin{bmatrix} 1+z & x-iy \\ x+iy & 1-z \end{bmatrix} \begin{bmatrix} e^{-i\theta_L} & 0 \\ 0 & e^{-i\theta_R} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1+z & (x-iy)e^{-i\varphi} \\ (x+iy)e^{i\varphi} & 1-z \end{bmatrix}$$

where, as before,  $\varphi \equiv \theta_R - \theta_L$  is the phase difference between the two neutron paths. The real and imaginary parts of the off-diagonal terms are:

$$\begin{aligned} (x+iy)e^{i\varphi} &= [x \cos(\varphi) - y \sin(\varphi)] + i[x \sin(\varphi) + y \cos(\varphi)] \\ (x-iy)e^{-i\varphi} &= [x \cos(\varphi) - y \sin(\varphi)] - i[x \sin(\varphi) + y \cos(\varphi)] \end{aligned}$$

As a real linear combination of  $\hat{1}$  and Pauli matrices, the transformed state  $\hat{\Phi}\hat{\rho}\hat{\Phi}^\dagger$  is:

$$\hat{\Phi}\hat{\rho}\hat{\Phi}^\dagger = \frac{1}{2} \left( \hat{1} + [x \cos(\varphi) - y \sin(\varphi)]\hat{\sigma}_x + [x \sin(\varphi) + y \cos(\varphi)]\hat{\sigma}_y + z\hat{\sigma}_z \right)$$

$\hat{\Phi}$  sends a state with Bloch vector  $(x, y, z)$  to  $(x \cos(\varphi) - y \sin(\varphi), x \sin(\varphi) + y \cos(\varphi), z)$ .

The matrix representing  $\hat{\Phi}$  in Pauli coordinates therefore must be:

$$\hat{\Phi} = \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) & 0 \\ \sin(\varphi) & \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

This matrix is a  $\varphi$  rotation about the  $z$ -axis. Figure 3.6 shows a visualization of acting  $\hat{S}_1$  on the excited state  $(0, 0, -1)$ , then acting  $\hat{\Phi}$  on the result with  $\varphi$  set to  $180^\circ$ .

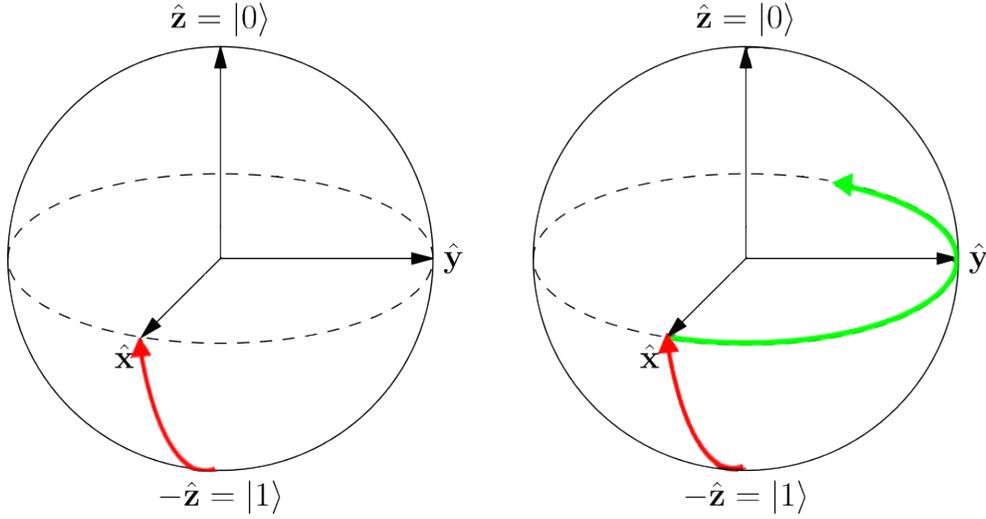


Figure 3.6: Action of the operator  $\hat{\Phi}\hat{S}_1$  on the excited state  $(0, 0, -1)$ .

Left:  $\hat{S}_1$  is a  $-90^\circ$  rotation about the  $y$ -axis. Right:  $\hat{\Phi}$  is a  $\varphi$  rotation about the  $z$ -axis.

(In this picture, the parameter  $\varphi$  was chosen to be  $180^\circ$ .)

In Bob's intended experiment, the neutron's initial state is  $|R\rangle$ , which has been renamed  $|1\rangle$ . The state immediately prior to detection is  $\hat{S}_1^\dagger \hat{\Phi} \hat{S}_1 |1\rangle$ . In Bloch-sphere language, Bob's experimental apparatus performs the composite operation "rotate  $-90^\circ$  about the  $y$ -axis, then rotate  $\varphi$  about the  $z$  axis, then rotate  $+90^\circ$  about the  $y$ -axis." In Pauli coordinates, the apparatus is represented by the operator  $\hat{G} \equiv \hat{S}_1^\dagger \hat{\Phi} \hat{S}_1$ :

$$\hat{G} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) & 0 \\ \sin(\varphi) & \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & \sin(\varphi) & \cos(\varphi) \end{bmatrix}$$

The initial state is  $(0, 0, -1) = -\hat{\mathbf{z}}$ , so the final state is  $-\hat{\mathbf{G}}\hat{\mathbf{z}}$ :

$$-\hat{\mathbf{G}}\hat{\mathbf{z}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & \sin(\varphi) & \cos(\varphi) \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ \sin(\varphi) \\ -\cos(\varphi) \end{bmatrix}$$

This result agrees with Bob's original prediction:

$$\hat{\rho} = \frac{1}{2} \left( \hat{1} + \sin(\varphi)\hat{\sigma}_y - \cos(\varphi)\hat{\sigma}_z \right) = \frac{1}{2} \begin{bmatrix} 1 - \cos(\varphi) & -i \sin(\varphi) \\ i \sin(\varphi) & 1 + \cos(\varphi) \end{bmatrix}$$

Define a **generalized Zeilinger-type apparatus** (GZA) as any device whose effect on a qubit state is represented in Pauli coordinates by a matrix of the form  $\hat{\mathbf{G}} \equiv \hat{\mathbf{S}}_1^\dagger \hat{\Phi} \hat{\mathbf{S}}_1$ . Roughly speaking, the action of the GZA is mathematically equivalent to slicing a neutron in half, sending its components on two distinct paths, then recombining the two parts coherently. Figure 3.7 shows a depiction of a half-cycle GZA, i.e.  $\hat{\mathbf{G}}$  with  $\varphi$  set to  $180^\circ$ .

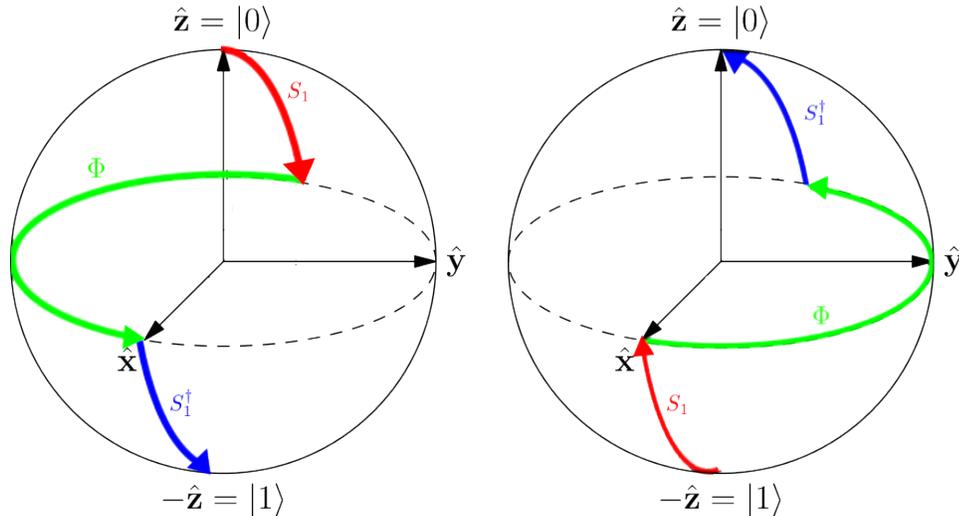


Figure 3.7: Action of the GZA on energy eigenstates.  
 Left: Half-cycle GZA ( $\varphi = 180^\circ$ ) applied to the ground state.  
 Right: The same operation applied to the excited state.

Bob's intended experiment is equivalent to acting a GZA on a qubit excited state  $|1\rangle$ , then measuring the qubit's energy. When Allyson's coin lands tails, she alters the GZA so that the third operation rotates  $\mathbf{r}$  in the opposite direction from what Bob intended.

To visualize the *1000 Cuts* experiment on the Bloch sphere, define an operator-valued random variable  $\hat{\mathbf{R}}_\varphi$  which is equivalent to  $\hat{\mathbf{G}}$  with the phase difference  $\varphi$  represented by a random variable rather than a constant. As before, let the initial state be  $(0, 0, -1)$ . If  $\varphi$  is

normally-distributed with mean  $\mu$  and variance  $\sigma^2$ , then the resulting mean state is:

$$E[-\hat{R}_\varphi \mathbf{z}] = - \int_{-\infty}^{\infty} [\hat{R}_\varphi \mathbf{z}] p(\varphi) d\varphi = \begin{bmatrix} 0 \\ \sin(\mu) e^{-\frac{1}{2}\sigma^2} \\ -\cos(\mu) e^{-\frac{1}{2}\sigma^2} \end{bmatrix}$$

Define a **randomized Zeilinger-type apparatus** (RZA) as any device whose effect on a qubit state is represented in Pauli coordinates by a matrix-valued random variable of the form  $\hat{R}_\varphi$ . The RZA is a “state mixer” which appears to erase quantum information by turning pure states into mixed states. More precisely, it performs a random rotation on the Bloch sphere to produce a random pure state whose expectation value is a mixed state. The RZA may be unpredictable, but it never destroys information.

These examples show the utility of Pauli and Bloch coordinates: qubit operations can be described in purely geometric terms without reference to neutrons, beamsplitters, or the physical details of a particular qubit design. (One might also consider an inverse question: is it possible to use neutrons and beamsplitters to simulate an arbitrary unitary operation on a finite-dimensional quantum state? According to [39], the answer is yes.)

### 3.3 Zech’s qubit

This thought experiment is loosely based on a colleague’s efforts to shield a Josephson-junction qubit from stray magnetic fields. Zech places a spin- $\frac{1}{2}$  particle inside a black box designed to isolate its contents from all interactions except a spatially-uniform magnetic field which can be very precisely controlled. Zech sets the controls to produce a field  $\mathbf{B}_0 = B_0 \mathbf{z}$  of constant magnitude in the  $z$  direction. The box is flawless except for one weakness: every day at exactly noon local time, the spin is exposed to a stray field which temporarily changes  $\mathbf{B}$  to  $B_0(\mathbf{z} + b\mathbf{x})$ .

Zech spends several weeks measuring the duration  $\tau$  and relative amplitude  $b$  of these stray fields. As far as he can tell, both numbers vary randomly from day to day with no predictable pattern. Today, he prepares the qubit in its ground state shortly before noon. Can he predict what the qubit’s state will be after the stray field has passed?

The Hamiltonian operator for this system is  $\hat{H} = -\gamma \mathbf{B} \cdot \hat{\mathbf{S}} = -\frac{1}{2}\gamma \hbar (B_x \hat{\sigma}_x + B_y \hat{\sigma}_y + B_z \hat{\sigma}_z)$ , where  $\gamma$  is the gyromagnetic ratio of the particle. Let  $\hat{H}_0$  denote the original Hamiltonian

and  $\hat{H}_\sharp$  the Hamiltonian whenever the stray field is active.

$$\hat{H}_0 = \frac{-\gamma\hbar B_0}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \hat{H}_\sharp = \frac{-\gamma\hbar B_0}{2} \begin{bmatrix} 1 & b \\ b & -1 \end{bmatrix}$$

For notational simplicity, define the natural angular frequency  $\omega_0 \equiv \gamma B_0$  and the ‘‘sharp’’ angular frequency  $\omega_\sharp \equiv \omega_0 \sqrt{1+b^2}$ . Eigenvalues of  $\hat{H}_0$  are then  $\pm \frac{1}{2}\hbar\omega_0$ , and eigenvalues of  $\hat{H}_1$  are  $\pm \frac{1}{2}\hbar\omega_\sharp$ . The Pauli coordinates of these operators are:

$$\mathbf{H}_0 = -\hbar\omega_0 \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T \quad \mathbf{H}_\sharp = -\hbar\omega_0 \begin{bmatrix} b & 0 & 1 \end{bmatrix}^T$$

The LvN equation is  $\hbar\dot{\mathbf{r}} = \mathbf{H} \times \mathbf{r}$ . When no stray field is present, it simplifies to:

$$\dot{\mathbf{r}} = -\omega_0 \mathbf{z} \times \mathbf{r} = \omega_0 \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \omega_0 \begin{bmatrix} y \\ -x \\ 0 \end{bmatrix}$$

The solution rotates  $\mathbf{r}$  about the  $z$ -axis with angular frequency  $-\omega_0$ .

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos(\omega_0 t) & \sin(\omega_0 t) & 0 \\ -\sin(\omega_0 t) & \cos(\omega_0 t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$$

The qubit’s initial state is the ground state  $(0, 0, 1)$ . Its state is unchanged until the stray field appears at time  $t = 0$ . While the stray field is present, the LvN equation is:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = -\omega_0 (b\mathbf{x} + \mathbf{z}) \times \mathbf{r} = \omega_0 \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & b \\ 0 & -b & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \omega_0 \begin{bmatrix} y \\ -x + bz \\ -by \end{bmatrix}$$

The solution rotates  $\mathbf{r}$  about the axis  $(b, 0, 1)$  with angular frequency  $-\omega_\sharp$ :<sup>7</sup>

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{1+b^2} \begin{bmatrix} b^2 + \cos(\omega_\sharp t) & \sqrt{1+b^2} \sin(\omega_\sharp t) & b - b \cos(\omega_\sharp t) \\ -\sqrt{1+b^2} \sin(\omega_\sharp t) & (1+b^2) \cos(\omega_\sharp t) & b\sqrt{1+b^2} \sin(\omega_\sharp t) \\ b - b \cos(\omega_\sharp t) & -b\sqrt{1+b^2} \sin(\omega_\sharp t) & 1 + b^2 \cos(\omega_\sharp t) \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$$

<sup>7</sup>This result was found by symbolically calculating  $\exp([\mathbf{H}_\sharp \times]t)$  in *Mathematica*.

After  $\tau$  seconds of exposure to the stray field, the qubit's state is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{1+b^2} \begin{bmatrix} b - b \cos(\omega_{\#}\tau) \\ b\sqrt{1+b^2} \sin(\omega_{\#}\tau) \\ 1 + b^2 \cos(\omega_{\#}\tau) \end{bmatrix}$$

After the stray field stops, the Hamiltonian reverts to  $\mathbf{H}_0$  and the qubit rotates about the  $z$ -axis for the next  $(t - \tau)$  seconds. The qubit's new state at time  $t > \tau$  is:

$$\begin{aligned} \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} &= \frac{1}{1+b^2} \begin{bmatrix} \cos(\omega_0[t-\tau]) & \sin(\omega_0[t-\tau]) & 0 \\ -\sin(\omega_0[t-\tau]) & \cos(\omega_0[t-\tau]) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} b(1 - \cos(\omega_{\#}\tau)) \\ b\sqrt{1+b^2} \sin(\omega_{\#}\tau) \\ 1 + b^2 \cos(\omega_{\#}\tau) \end{bmatrix} \\ &= \frac{1}{1+b^2} \begin{bmatrix} b \cos(\omega_0[t-\tau])(1 - \cos(\omega_{\#}\tau)) + b\sqrt{1+b^2} \sin(\omega_0[t-\tau]) \sin(\omega_{\#}\tau) \\ -b \sin(\omega_0[t-\tau])(1 - \cos(\omega_{\#}\tau)) + b\sqrt{1+b^2} \cos(\omega_0[t-\tau]) \sin(\omega_{\#}\tau) \\ 1 + b^2 \cos(\omega_{\#}\tau) \end{bmatrix} \end{aligned}$$

If Zech measures the qubit's energy at time  $t > \tau$ , then the Born rule predicts he will find the higher energy eigenvalue with probability  $P(\text{excited}|\mathbf{r}) = \frac{1}{2}(1 - z)$ .

$$P(\text{excited}|\mathbf{r}) = \frac{1}{2}(1 - z) = \frac{1}{2} - \frac{1 + b^2 \cos(\omega_{\#}\tau)}{2(1 + b^2)} = \frac{b^2}{2(1 + b^2)} [1 - \cos(\omega_{\#}\tau)]$$

Figure 3.8 shows  $P(\text{excited}|\mathbf{r})$  for  $\tau \in [0, 4\pi]$  and  $b \in [0, 2]$ . For weak stray fields with  $|b| \ll 1$ , the qubit is barely affected. As  $|b|$  increases, oscillations become larger and faster.

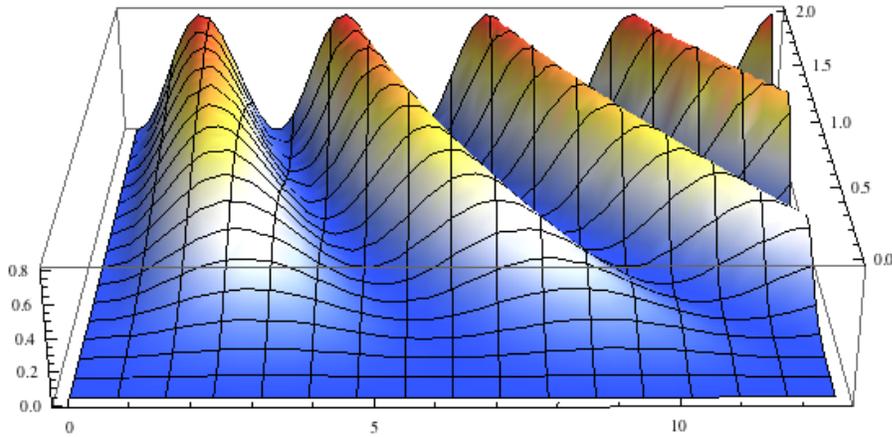


Figure 3.8:  $P(\text{excited})$  as a function of stray-field amplitude  $b$  and duration  $\tau$ . Left-right axis is  $\omega_0\tau \in [0, 4\pi]$ . Front-back axis is  $b \in [0, 4]$ . Vertical axis is  $P(\text{excited}|\mathbf{r})$ .

If Zech does not precisely know  $\tau$  and  $b$ , then he does not know what state is actually present at  $t > 0$ , but can make an “educated guess” by calculating a mean state. For calculational simplicity, suppose he represents  $\tau$  and  $b$  as uniformly-distributed random variables on the intervals  $[0, \mathcal{T}]$  and  $[-\mathcal{B}, \mathcal{B}]$ , respectively. (A negative value for  $b$  means the stray field points in the  $-x$  direction.) The mean state is the expectation value of the true state averaged over all possible  $b$  values and all possible  $\tau$  values:

$$\bar{\mathbf{r}}(t) = \int_0^{\mathcal{T}} \int_{-\mathcal{B}}^{\mathcal{B}} \mathbf{r}_t(b, \tau) p(b, \tau) db d\tau = \frac{1}{2\mathcal{B}\mathcal{T}} \int_0^{\mathcal{T}} \int_{-\mathcal{B}}^{\mathcal{B}} \mathbf{r}_t(b, \tau) db d\tau$$

The mean state coordinates  $\bar{x}, \bar{y}$  are zero because the integrands are odd functions of  $b$ :

$$\begin{aligned} x_t(b, \tau) &= \frac{b \cos(\omega_0[t - \tau])(1 - \cos(\omega_{\sharp}\tau)) + b\sqrt{1 + b^2} \sin(\omega_0[t - \tau]) \sin(\omega_{\sharp}\tau)}{1 + b^2} \\ y_t(b, \tau) &= \frac{-b \sin(\omega_0[t - \tau])(1 - \cos(\omega_{\sharp}\tau)) + b\sqrt{1 + b^2} \cos(\omega_0[t - \tau]) \sin(\omega_{\sharp}\tau)}{1 + b^2} \\ \Rightarrow \int_{-\mathcal{B}}^{\mathcal{B}} x_t(b, \tau) db &= 0 \quad \text{and} \quad \int_{-\mathcal{B}}^{\mathcal{B}} y_t(b, \tau) db = 0 \end{aligned}$$

Note that  $\omega_{\sharp} = \omega_0\sqrt{1 + b^2}$  depends on  $b$ . The  $\bar{z}$  coordinate is not as easy:

$$\begin{aligned} \bar{z}(t) &= \frac{1}{2\mathcal{B}\mathcal{T}} \int_0^{\mathcal{T}} \int_{-\mathcal{B}}^{\mathcal{B}} \frac{1 + b^2 \cos(\omega_0\tau\sqrt{1 + b^2})}{1 + b^2} db d\tau \\ &= \frac{1}{2\mathcal{B}\mathcal{T}} \int_{-\mathcal{B}}^{\mathcal{B}} \frac{\mathcal{T}}{1 + b^2} + \frac{b^2 \sin(\omega_0\mathcal{T}\sqrt{1 + b^2})}{(1 + b^2)^{3/2}} db \\ &= \frac{\arctan(\mathcal{B})}{\mathcal{B}} + \frac{1}{2\mathcal{B}\mathcal{T}} \int_{-\mathcal{B}}^{\mathcal{B}} \frac{b^2 \sin(\omega_0\mathcal{T}\sqrt{1 + b^2})}{(1 + b^2)^{3/2}} db \end{aligned}$$

In the long-duration limit  $\mathcal{T} \rightarrow \infty$ , the remaining integral vanishes.

$$\lim_{\mathcal{T} \rightarrow \infty} \bar{z}(t) = \frac{\arctan(\mathcal{B})}{\mathcal{B}}$$

In this limit, Zech’s mean state  $\bar{\rho}_Z$  resembles the canonical qubit density matrix  $\bar{\rho}_C$ .

$$\bar{\rho}_Z = \frac{1}{2} \begin{bmatrix} 1 + \frac{1}{\mathcal{B}} \arctan(\mathcal{B}) & 0 \\ 0 & 1 - \frac{1}{\mathcal{B}} \arctan(\mathcal{B}) \end{bmatrix} \quad \bar{\rho}_C = \frac{1}{2} \begin{bmatrix} 1 + \tanh\left(\frac{\hbar\omega_0}{2kT}\right) & 0 \\ 0 & 1 - \tanh\left(\frac{\hbar\omega_0}{2kT}\right) \end{bmatrix}$$

To see the resemblance, define the **pseudotemperature**  $kT$  of Zech’s mean state as the expected relative stray field magnitude times the qubit’s energy gap  $\hbar\omega_0$ :

$$kT \equiv \hbar\omega_0 E[|b|] = \frac{\hbar\omega_0}{2\mathcal{B}} \int_{-\mathcal{B}}^{\mathcal{B}} |b| db = \frac{\hbar\omega_0}{\mathcal{B}} \int_0^{\mathcal{B}} b db = \frac{1}{2} \hbar\omega_0 \mathcal{B}$$

Substituting this quantity for  $kT$  into the canonical density matrix gives:

$$P(\text{excited}) = \text{logistic}\left(-\frac{2}{\mathcal{B}}\right) = \frac{1}{2}\left[1 - \tanh\left(\frac{1}{\mathcal{B}}\right)\right]$$

This result does *not* exactly agree with Zech’s mean state, but it is qualitatively similar. Figure 3.9 shows two plots of  $P(\text{excited})$  as a function of  $\mathcal{B}$ : one for Zech’s mean state, and another for the canonical density matrix with the substitution  $kT \rightarrow \frac{1}{2}\hbar\omega_0\mathcal{B}$ .

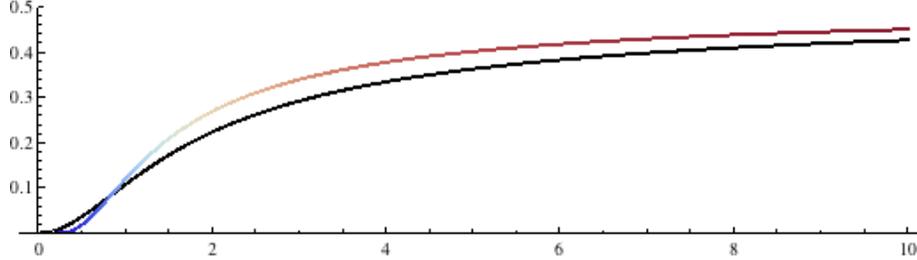


Figure 3.9:  $P(\text{excited})$  as a function of maximum stray-field strength  $\mathcal{B}$ . Black: Zech’s  $\bar{\rho}_Z$  with  $\mathcal{T} \rightarrow \infty$ . Color: Canonical  $\bar{\rho}_C$  with  $kT = \frac{1}{2}\hbar\omega_0\mathcal{B}$ .

The analogy between temperature and pseudotemperature should not be taken too seriously. The quantity  $\mathcal{B}$  describes how violently Zech’s experiment is scrambled by environmental interactions, which roughly agrees with physicists’ intuition about temperature. But the “pseudo-” prefix is used to stress that Zech’s  $\mathcal{B}$  is *subjective*: it depends on how precisely he can predict the stray field. By contrast, temperature of a canonical qubit density matrix is defined by an objective procedure:  $\beta \equiv S'(U)$ , where  $S(U)$  is the maximum von Neumann entropy of a qubit mean state with the constraint  $\langle H \rangle = U$ .

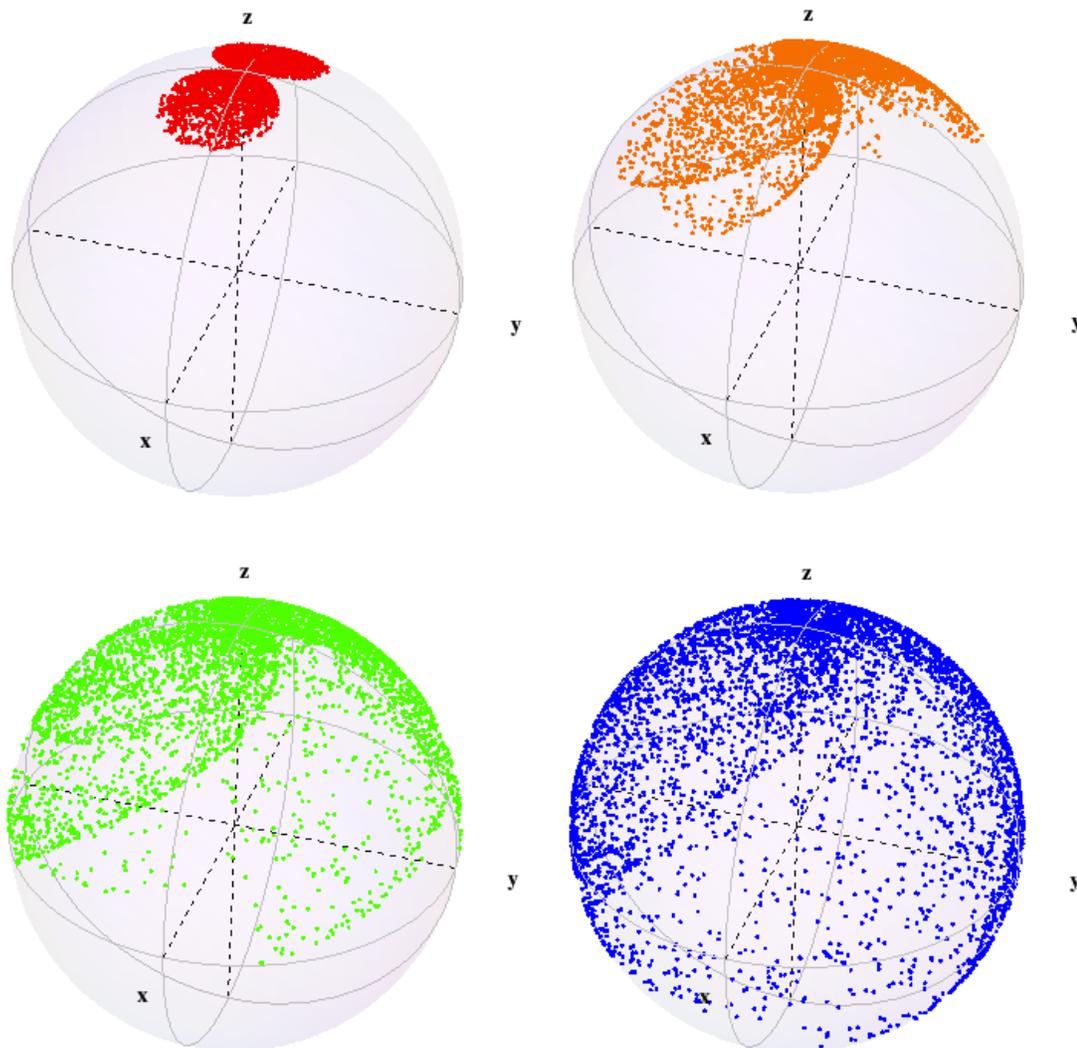
In this thought experiment, the Loschmidt paradox is: “How did Zech’s qubit gain entropy if it evolved unitarily?” The paradox can be resolved by replacing the phrase “entropy of the qubit” with the phrase “entropy of the average taken over a distribution of possible states.” In other words: *Zech’s mean state*  $\bar{\mathbf{r}}(t)$  gained entropy and the true state did not.

The qubit’s true state has been encrypted but not irreversibly destroyed. In this case, the password is knowledge of the stray field’s duration  $\tau$  and relative amplitude  $b$ . If Zech has somehow precisely recorded these quantities, then he can “rewind” the qubit using a procedure similar to the *spin-echo* methods explained in [40]. First he applies a field  $-\mathbf{B}_0$  for  $(t - \tau)$  seconds, then he reverses the effect of the stray field by applying  $-\mathbf{B}_\ddagger$  for  $\tau$  seconds. The qubit’s Bloch vector  $\mathbf{r}$  then follows its previous trajectory in reverse.

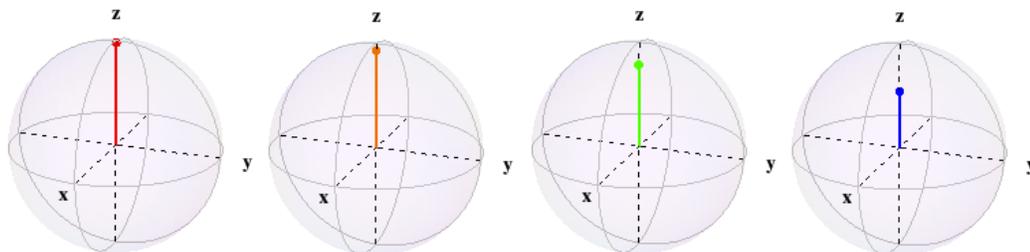
Figure 3.10 shows four Monte Carlo simulations of Zech’s qubit. If the stray field was known to be weak and short-lived, then the true state is probably close to the state Zech intended. In the opposite limit,  $\bar{\mathbf{r}} \approx \mathbf{0}$  and Zech has no idea what state is actually present.

Figure 3.10: Monte Carlo simulations of Zech's qubit at time  $\omega_0 t = 10\pi$ .

5000 pseudorandom states are shown for stray-field duration  $\omega_0 \tau \in [0, 4\pi]$  and relative amplitude  $b \in [-\mathcal{B}, \mathcal{B}]$ . Colors indicate maximum stray-field magnitude  $\mathcal{B} = \{ \frac{1}{4}, \frac{1}{2}, 1, 2 \}$



Estimated mean states  $\bar{\mathbf{r}}$  are found by averaging all samples of each color.



## 4 STOCHASTIC QUBITS

This chapter attempts to focus on *using* results of stochastic calculus rather than *proving* them. For readers who are not content to “shut up and calculate,” Appendix A provides definitions and references to more rigorous mathematical explanations.

Drunk models do not attempt to describe the causes of experimental errors in detail. In the spirit of Wigner’s random matrix models, environmental interactions are “assumed to be so complicated that statistical considerations can be applied to them.” [41] Laboratories are modeled as gambling machines whose statistical properties can be defined, but whose details are too unpredictable to be represented precisely. The strengths and weaknesses of environment-as-casino theories were summarized by Tanimura and Kubo:

The underlying stochastic process is merely a model appropriate for the problem rather than something to be derived from microscopic considerations. This is regarded as an advantage since it can cover a wide category of physical cases from a unified point of view. Furthermore the calculations may be carried out by non-perturbative methods.[42]

Each model in this chapter begins with a heuristic time-evolution equation of the form:

$$\frac{d\mathbf{r}_t}{dt} = -\mathbf{B}_t \times \mathbf{r}_t \quad \mathbf{B}_t = \text{the } \mathbf{B}(t) \text{ a scientist intended} + \text{random noise}$$

Here the true state  $\mathbf{r}_t$  and true Hamiltonian  $\mathbf{B}_t$  are  $\mathbb{R}^3$ -valued stochastic processes. To produce useful predictions, this vague statement requires a mathematical representation of “random noise.” A slightly less ambiguous statement is this Langevin-type equation:

$$\frac{d\mathbf{r}_t}{dt} = -\mathbf{B}_t \times \mathbf{r}_t \quad \mathbf{B}_t = \boldsymbol{\mu}(t, \mathbf{r}_t) + \hat{\Sigma}(t, \mathbf{r}_t) \frac{d\mathbf{W}_t}{dt} \quad (\text{Pre})$$

The **mean field**  $\boldsymbol{\mu}(t, \mathbf{r}_t)$  is a real 3D vector, the **volatility matrix**  $\hat{\Sigma}(t, \mathbf{r}_t)$  is a real  $3 \times 3$  matrix,  $\mathbf{W}_t$  is a real 3D vector whose components are independent Wiener processes, and the symbol  $d\mathbf{W}_t/dt$  is a poorly-defined avatar for Gaussian white noise. This symbol is a useful heuristic, but it is ambiguous because  $\mathbf{W}_t$  is almost nowhere differentiable. The expression (Pre) is what van Kampen called a “pre-equation” which “is really a meaningless string of symbols” until one specifies which stochastic calculus to use.[43]

Itô and Stratonovich-Fisk (SF) calculus provide two distinct formalisms for defining stochastic differential equations (SDEs). A famous peculiarity of Itô calculus is that the Chain Rule from ordinary calculus must be replaced with *Itô's Lemma*, which makes coordinate-free formulas potentially confusing. The SF Chain Rule resembles ordinary calculus, but expectation values of SF integrals can be much more difficult to calculate than their Itô counterparts. Appendix A briefly reviews the Itô-SF relationship. For thorough explanations, see van Kampen, Sussmann, and their cited references.[43][44][45]

This chapter uses *both* calculi: a noise model is used to derive an SF SDE which is then converted to an equivalent Itô SDE in order to more easily calculate expectation values. An equation of the form (Pre) must be considered an SF SDE or else the possible states violate the unitary-evolution assumption  $|\mathbf{r}_t\rangle = |\mathbf{r}_0\rangle$ .<sup>1</sup> An equivalent Itô SDE can be found by using the *Wong-Zakai correction* formula in Appendix A. The equivalent Itô SDE predicts the same measurement probabilities, but it generally cannot be written in the form (Pre).

All models in this chapter follow the same algorithm with different choices of  $\boldsymbol{\mu}$  and  $\hat{\Sigma}$ :

1. Define a heuristic pre-equation  $\frac{d}{dt}\mathbf{r}_t = -\mathbf{B}_t \times \mathbf{r}_t$  with  $\mathbf{B}_t = \boldsymbol{\mu}(t, \mathbf{r}_t) + \hat{\Sigma}(t, \mathbf{r}_t)\frac{d\mathbf{W}}{dt}$ .
2. “Multiply by  $dt$ ” and proclaim the result to be a Stratonovich-Fisk SDE.
3. If  $\boldsymbol{\mu}$  and  $\hat{\Sigma}$  do not depend on  $\mathbf{r}_t$ , then solve the *drunken master equation* in Section 4.1 to find the model’s mean state  $\bar{\mathbf{r}}(t)$ .
4. Use the numerical methods in Chapter 5 to simulate experiments.

The use of Gaussian white noise models here is motivated primarily by mathematical convenience. A partial justification is given by the Central Limit Theorem: if many independent errors occur very rapidly, then their sum tends to a Gaussian distribution.<sup>2</sup> However, these models might not be appropriate for all experiments. For example, there is evidence that **pink noise** with a  $1/f$  power spectral density affects superconducting qubits.[47]

Stochastic drunk models have many historical precedents. Kubo and Hashitsume derived the NMR Bloch equation using very similar mathematics in 1970.[48] Fox used Gaussian white noise models to find a quantum version of Boltzmann’s H-theorem and another Bloch-equation derivation.[49][50][51] Gorini, Kossakowski, and Sudarshan developed similar models for noisy  $N$ -level quantum systems.[52][53] In the 1980’s, Tanimura and Kubo considered quantum systems with stochastic Gaussian-Markovian environments.[42] More recently, Stockburger and Mak used stochastic Liouville-von Neumann equations to model dissipative quantum systems.[54][55] Tanimura’s 2005 review references “over half a century” of work on stochastic quantum evolution.[56]

<sup>1</sup>For proof, use the SF Chain Rule from Appendix A or see e.g. [46].

<sup>2</sup>The errors need not be identically-distributed over time so long as the *Lyapunov criterion* is satisfied.

## 4.1 The Drunken Master Equation

A **drunken master equation** is an ordinary differential equation (ODE) for the mean state of a drunk model. A drunk model is **linear** if the mean field  $\boldsymbol{\mu}$  and volatility matrix  $\hat{\Sigma}$  do not depend on  $\mathbf{r}_t$ . Each linear model has its own linear drunken master equation. Solving a linear ODE is typically much easier than solving an SDE or a partial differential equation for the true state's probability density  $p(\mathbf{r}, t)d^3\mathbf{r}$ .

Following van Kampen's advice in [43], promote pre-equation (Pre) to a legitimate SDE by "multiplying by  $dt$ " and choosing Stratonovich-Fisk stochastic calculus:

$$d\mathbf{r}_t = -\left(\boldsymbol{\mu}(t, \mathbf{r}_t)dt + \hat{\Sigma}(t, \mathbf{r}_t) \circ d\mathbf{W}\right) \times \mathbf{r}_t \quad (4.1)$$

The mean field  $\boldsymbol{\mu}$  and volatility matrix  $\hat{\Sigma}$  are deterministic functions. The  $\circ d\mathbf{W}$  denotes "use SF calculus," while  $d\mathbf{W}$  without the  $\circ$  means "use Itô calculus."

Some calculations are easier if the components of  $d\mathbf{W}$  are written separately as  $dW^1, dW^2, dW^3$ . (The superscripts are indices, not exponents.) Define three **volatility vectors**  $\{\boldsymbol{\nu}_m(\mathbf{r}_t, t)\}$ , each of which is the  $m$ th column of  $\hat{\Sigma}(\mathbf{r}_t, t)$ . For notational clarity, temporarily hide any  $t$  and/or  $\mathbf{r}_t$  dependence of  $\boldsymbol{\nu}_1, \boldsymbol{\nu}_2, \boldsymbol{\nu}_3$  and  $\boldsymbol{\mu}$ . Equation (4.1) is then:

$$d\mathbf{r}_t = -\left(\boldsymbol{\mu} \times \mathbf{r}_t\right)dt - \sum_{m=1}^3 \left(\boldsymbol{\nu}_m \times \mathbf{r}_t\right) \circ dW^m \quad (\text{SF SDE})$$

The Wong-Zakai correction from Appendix A converts (SF SDE) into an Itô SDE:

$$d\mathbf{r}_t = -\left(\boldsymbol{\mu} \times \mathbf{r}_t\right)dt - \sum_{m=1}^3 \left(\boldsymbol{\nu}_m \times \mathbf{r}_t\right)dW^m + \frac{1}{2} \sum_{m=1}^3 \hat{D}(\boldsymbol{\nu}_m \times \mathbf{r}_t)(\boldsymbol{\nu}_m \times \mathbf{r}_t) dt \quad (\text{Itô SDE})$$

Here  $\hat{D}(\boldsymbol{\nu}_m \times \mathbf{r}_t)$  denotes the a Jacobian matrix of first-partial derivatives of  $(\boldsymbol{\nu}_m \times \mathbf{r}_t)$  with respect to  $x_t, y_t, z_t$ . This matrix is often easier to calculate if the cross-product is written as a linear operator. For any vector  $\mathbf{a} \in \mathbb{R}^3$ , let  $[\mathbf{a} \times]$  denote the matrix  $a_x \hat{I} + a_y \hat{J} + a_z \hat{K}$ :

$$[\mathbf{a} \times] \equiv a_x \hat{I} + a_y \hat{J} + a_z \hat{K} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix}$$

Then  $[\mathbf{a} \times] \mathbf{r}$  is the same vector as  $\mathbf{a} \times \mathbf{r}$ . In this notation, (Itô SDE) is:

$$d\mathbf{r}_t = -[\boldsymbol{\mu} \times] \mathbf{r}_t dt - \sum_{m=1}^3 [\boldsymbol{\nu}_m \times] \mathbf{r}_t dW^m + \frac{1}{2} \sum_{m=1}^3 \hat{D}([\boldsymbol{\nu}_m \times] \mathbf{r}_t) [\boldsymbol{\nu}_m \times] \mathbf{r}_t dt$$

The extra terms in (Itô SDE) are a nuisance, but the *martingale property* of Itô integrals greatly simplifies calculation of the mean state  $\bar{\mathbf{r}}(t) \equiv E[\mathbf{r}_t]$ . For the present purposes, the martingale property means the  $dW^m$  terms in (Itô SDE) contribute nothing to  $E[\mathbf{r}_t]$ .<sup>3</sup>

For linear models, the diffusion terms  $\{\boldsymbol{\nu}_m\}$  have no  $\mathbf{r}_t$  dependence and  $\hat{D}$  simplifies:

$$\boxed{d\mathbf{r}_t = -[\boldsymbol{\mu} \times] \mathbf{r}_t dt - \sum_{m=1}^3 [\boldsymbol{\nu}_m \times] \mathbf{r}_t dW^m + \frac{1}{2} \sum_{m=1}^3 [\boldsymbol{\nu}_m \times]^2 \mathbf{r}_t dt} \quad (\text{Linear Itô SDE})$$

Take the expectation value of both sides of (Linear Itô SDE) and erase the  $dW^m$  terms:<sup>4</sup>

$$dE[\mathbf{r}_t] = E \left[ -[\boldsymbol{\mu} \times] \mathbf{r}_t + \frac{1}{2} \sum_{m=1}^3 [\boldsymbol{\nu}_m \times]^2 \mathbf{r}_t \right] dt$$

Linearity of expectation values then implies:

$$dE[\mathbf{r}_t] = \left[ -[\boldsymbol{\mu} \times] E[\mathbf{r}_t] + \frac{1}{2} \sum_{m=1}^3 [\boldsymbol{\nu}_m \times]^2 E[\mathbf{r}_t] \right] dt$$

In general, one does not simply assume  $E[\mathbf{f}(\mathbf{r}_t)] = \mathbf{f}(E[\mathbf{r}_t])$ , as that would violate Jensen's inequality unless  $\mathbf{f}$  is affine. But for linear models,  $[\boldsymbol{\mu} \times] \mathbf{r}_t$  and  $[\boldsymbol{\nu}_m \times]^2 \mathbf{r}_t$  are linear functions of  $\mathbf{r}_t$  and the assumption is valid. Using  $\bar{\mathbf{r}}(t) \equiv E[\mathbf{r}_t]$ , the previous equation is:

$$\boxed{\frac{d}{dt} \bar{\mathbf{r}} = -[\boldsymbol{\mu} \times] \bar{\mathbf{r}} + \frac{1}{2} \sum_{m=1}^3 [\boldsymbol{\nu}_m \times]^2 \bar{\mathbf{r}}} \quad (\text{Master})$$

(Master) is valid whether or not the mean field  $\boldsymbol{\mu}$  and volatility  $\hat{\Sigma}$  depend on time. But if either depends on  $\mathbf{r}_t$ , then (Master) is *not* valid! For nonlinear models,  $\hat{D}(\boldsymbol{\nu}_m \times \mathbf{r}_t)$  is more complicated and Jensen's inequality ruins the last step in the derivation.

For practical applications, it may be useful to write (Master) in various other notations.

<sup>3</sup>Here  $E[\mathbf{r}_t]$  is short notation for  $E[\mathbf{r}_t | \mathcal{F}_t]$ , the expectation value of  $\mathbf{r}_t$  over its *natural filtration*  $\mathcal{F}_t$ . For thorough explanations of martingales and filtrations, see e.g. [57][58][59].

<sup>4</sup>Readers skeptical about this reckless use of infinitesimals  $dE[\mathbf{r}_t]$  and  $dt$  can find a more careful derivation with the same conclusion for linear Itô SDEs in [59].

For those who prefer traditional cross-product notation, the master equation is:

$$\frac{d}{dt}\bar{\mathbf{r}} = -\boldsymbol{\mu} \times \bar{\mathbf{r}} + \frac{1}{2} \sum_{m=1}^3 \boldsymbol{\nu}_m \times \boldsymbol{\nu}_m \times \bar{\mathbf{r}}$$

In terms of the **infinitesimal covariance matrix**  $\hat{\Sigma}^T \hat{\Sigma}$ , it is:

$$\frac{d}{dt}\bar{\mathbf{r}} = -\boldsymbol{\mu} \times \bar{\mathbf{r}} + \frac{1}{2} \left( \hat{\Sigma}^T \hat{\Sigma} - \text{Tr}[\hat{\Sigma}^T \hat{\Sigma}] \hat{\mathbf{1}} \right) \bar{\mathbf{r}}$$

In terms of rotation generators  $\hat{I}, \hat{J}, \hat{K}$ , it is:

$$\frac{d}{dt}\bar{\mathbf{r}} = - \left[ \mu_x \hat{I} + \mu_y \hat{J} + \mu_z \hat{K} \right] \bar{\mathbf{r}} + \frac{1}{2} \sum_{m=1}^3 \left[ \nu_{1m} \hat{I} + \nu_{2m} \hat{J} + \nu_{3m} \hat{K} \right]^2 \bar{\mathbf{r}}$$

Here  $\nu_{nm}$  is the  $n$ th component of the  $m$ th column of  $\hat{\Sigma}$ . In component notation:

$$\begin{aligned} \frac{d}{dt} \begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix} &= \begin{bmatrix} 0 & \mu_z & -\mu_y \\ -\mu_z & 0 & \mu_x \\ \mu_y & -\mu_x & 0 \end{bmatrix} \begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix} \\ &+ \frac{1}{2} \sum_{m=1}^3 \begin{bmatrix} -(\nu_{2m}^2 + \nu_{3m}^2) & \nu_{1m}\nu_{2m} & \nu_{1m}\nu_{3m} \\ \nu_{1m}\nu_{2m} & -(\nu_{3m}^2 + \nu_{1m}^2) & \nu_{2m}\nu_{3m} \\ \nu_{1m}\nu_{3m} & \nu_{2m}\nu_{3m} & -(\nu_{1m}^2 + \nu_{2m}^2) \end{bmatrix} \begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix} \end{aligned}$$

It is often convenient to consolidate (Master) into one **generator equation**:

$$\frac{d}{dt}\bar{\mathbf{r}}(t) = \hat{G}(t)\bar{\mathbf{r}}(t) \quad \hat{G}(t) \equiv -[\boldsymbol{\mu}(t) \times] + \frac{1}{2} \sum_{m=1}^3 [\boldsymbol{\nu}_m(t) \times]^2$$

If  $[\hat{G}(t), \hat{G}(s)] = 0$  for all  $s, t > 0$ , then the solution is a matrix exponential:

$$\bar{\mathbf{r}}(t) = \exp \left[ \int_0^t \hat{G}(s) ds \right] \bar{\mathbf{r}}(0)$$

Unfortunately, the assumption  $[\hat{G}(t), \hat{G}(s)] = 0$  is often wrong unless  $\boldsymbol{\mu}$  and  $\hat{\Sigma}$  are both constant in time.<sup>5</sup> Numerical methods for solving (Master) are explained in Chapter 5.

Suppose a scientist does not precisely know the initial state  $\mathbf{r}_0$  of a qubit, but he or she can construct an initial mean state  $\bar{\mathbf{r}}(0) \equiv E[\mathbf{r}_0 | \mathcal{F}_0]$ . If the initial state is known with certainty, then the initial filtration  $\mathcal{F}_0$  includes only a single element, and  $\bar{\mathbf{r}}(t)$  is an average over all

<sup>5</sup>The matrices  $\hat{I}, \hat{J}, \hat{K}$  are generators of the rotation group  $SO(3)$ , which is not commutative.

possible trajectories of  $\mathbf{r}_t$  with starting point  $\mathbf{r}_0$ . If many initial states are possible, then  $\bar{\mathbf{r}}(t) \equiv E[\mathbf{r}_t | \mathcal{F}_t]$  is an average over all trajectories *including all possible starting points*.

For conceptual clarity, assume the qubit's initial state is a discrete random variable: it might be one of several possible pure states  $\{\mathbf{r}_0^i\}$  with probabilities  $\{p_i\}$ . Let  $\bar{\mathbf{r}}_i(t)$  denote the expectation value, given fixed initial condition  $\mathbf{r}_0^i$ , of  $\mathbf{r}_t$  over all possible Hamiltonians. This  $\bar{\mathbf{r}}_i(t)$  is found by solving the master equation for initial condition  $\mathbf{r}_0^i$ . The Law of Total Probability requires the mean state at time  $t$  to be:

$$\bar{\mathbf{r}}(t) = \sum_i P(\mathbf{r}_0^i) E[\mathbf{r}(t) | \mathbf{r}_0^i] = \sum_i p_i \bar{\mathbf{r}}_i(t)$$

The master equation is a linear ODE, so each  $\bar{\mathbf{r}}_i(t)$  is a linear transformation of  $\mathbf{r}_{0i}$ . Let  $\hat{T}$  denote the operator which sends initial state  $\mathbf{r}_{0i}$  to its future mean state  $\bar{\mathbf{r}}_i(t)$ . Then

$$\bar{\mathbf{r}}(t) = \sum_i p_i \bar{\mathbf{r}}_i(t) = \sum_i p_i \hat{T} \mathbf{r}_0^i = \hat{T} \sum_i p_i \mathbf{r}_0^i = \hat{T} \bar{\mathbf{r}}(0)$$

Linear master equations make calculation of conditional mean states  $\bar{\mathbf{r}}_i$  unnecessary. Given a random initial state, one can calculate an initial mixed state  $\bar{\mathbf{r}}(0) = \sum p_i \mathbf{r}_0^i$ , then predict future mean states  $\bar{\mathbf{r}}(t)$  by solving the master equation using  $\bar{\mathbf{r}}(0)$  as the initial condition. (Like the master equation itself, this tactic is *not* valid for nonlinear models.)

## 4.2 The Second Law of Drunk Dynamics

Possible states of a drunk qubit tend to spread out as they wander randomly around the Bloch sphere. Though each possible state remains pure, the mean state  $\bar{\mathbf{r}}(t)$  tends to drift inward as  $t$  increases. For linear drunk models, this tendency can be made precise by a theorem which is here called the **Second Law of Drunk Dynamics**:

For any linear drunk qubit model with a nonsingular volatility matrix, von Neumann entropy of the mean state always increases.

The name is a somewhat facetious reference to the Second Law of Thermodynamics. There are no Zeroth, First, or Third Laws of Drunk Dynamics, and the proof neither requires nor uses any thermodynamic assumptions. In fact, it illustrates a disagreement between linear drunk models and qubit thermodynamics: the mean state of a linear drunk model *cannot* have a finite-temperature steady-state solution unless  $\hat{\Sigma}$  is singular. For nonsingular  $\hat{\Sigma}$ , the only possible steady-solution is  $\bar{\mathbf{r}}(\infty) \rightarrow \mathbf{0}$ , which is the canonical thermal mixed state only in the limits  $T \rightarrow \pm\infty$  or equivalently,  $\beta \rightarrow 0$ .

A geometric proof of the Second Law begins with Equation (Master):

$$\frac{d}{dt}\bar{\mathbf{r}} = \left( -[\boldsymbol{\mu} \times] + \frac{1}{2} \sum_{m=1}^3 [\boldsymbol{\nu}_m \times][\boldsymbol{\nu}_m \times] \right) \bar{\mathbf{r}}$$

As before,  $\boldsymbol{\nu}_m$  is the  $m$ th column of the volatility matrix  $\hat{\Sigma}$ , and any time dependence has been hidden for notational clarity. Von Neumann entropy  $S(\bar{\mathbf{r}})$  increases monotonically as a mean state's radius  $|\bar{\mathbf{r}}|$  decreases. With this in mind, use (Master) to find  $\frac{d}{dt}|\bar{\mathbf{r}}|^2$ :

$$\frac{d}{dt}|\bar{\mathbf{r}}|^2 = 2\bar{\mathbf{r}} \cdot \frac{d}{dt}\bar{\mathbf{r}} = -2\bar{\mathbf{r}} \cdot [\boldsymbol{\mu} \times]\bar{\mathbf{r}} + \bar{\mathbf{r}} \cdot \sum_{m=1}^3 [\boldsymbol{\nu}_m \times]^2 \bar{\mathbf{r}} = 0 + \sum_{m=1}^3 \bar{\mathbf{r}} \cdot [\boldsymbol{\nu}_m \times]^2 \bar{\mathbf{r}}$$

The first term is zero because  $\bar{\mathbf{r}} \cdot (\boldsymbol{\mu} \times \bar{\mathbf{r}}) = 0$  regardless of what  $\boldsymbol{\mu}$  is. Using  $[\mathbf{a} \times][\mathbf{b} \times]\mathbf{c} = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b})$ , the surviving term is:

$$\frac{d}{dt}|\bar{\mathbf{r}}|^2 = \sum_{m=1}^3 (\bar{\mathbf{r}} \cdot \boldsymbol{\nu}_m)^2 - |\bar{\mathbf{r}}|^2 |\boldsymbol{\nu}_m|^2 \quad (4.2)$$

For each nonzero column  $\boldsymbol{\nu}_m$ , let  $\varphi_m$  denote the smallest angle between  $\boldsymbol{\nu}_m$  and  $\bar{\mathbf{r}}$ . Use the Euclidean dot-product rule  $(\bar{\mathbf{r}} \cdot \boldsymbol{\nu}_m)^2 = |\bar{\mathbf{r}}|^2 |\boldsymbol{\nu}_m|^2 \cos^2(\varphi_m)$ :

$$\frac{d}{dt}|\bar{\mathbf{r}}|^2 = \sum_{m=1}^3 |\bar{\mathbf{r}}|^2 |\boldsymbol{\nu}_m|^2 [\cos^2(\varphi_m) - 1]$$

The  $m$ th term is zero if  $\bar{\mathbf{r}} = \mathbf{0}$  or  $\boldsymbol{\nu}_m = \mathbf{0}$  or  $\cos(\varphi_m) = \pm 1$ ; otherwise it is negative. The right side of (4.2) therefore cannot be positive, and it must be negative unless all three terms are zero. If all three terms are zero, then at least one of the following must be true:

1.  $\bar{\mathbf{r}} = \mathbf{0}$ .
2. All three  $\{\boldsymbol{\nu}_m\}$  have norm 0, which means  $\hat{\Sigma}$  is the zero matrix.
3. Each of the angles  $\varphi_1, \varphi_2, \varphi_3$  is either 0 or  $\pi$  radians.

Case 3 requires the  $\{\boldsymbol{\nu}_m\}$  to be linearly dependent. These vectors are columns of  $\hat{\Sigma}$ , so they cannot be linearly dependent unless  $\hat{\Sigma}$  is singular. A consequence of these results is:

$$\text{Det}[\hat{\Sigma}] \neq 0 \quad \text{and} \quad \bar{\mathbf{r}} \neq \mathbf{0} \quad \Rightarrow \quad \frac{d}{dt}|\bar{\mathbf{r}}|^2 < 0$$

Stronger statements are also possible. If  $\bar{\mathbf{r}} \neq \mathbf{0}$  and  $\hat{\Sigma} \neq \hat{0}$ , then  $\frac{d}{dt}|\bar{\mathbf{r}}|^2$  is strictly negative unless  $\bar{\mathbf{r}}$  is parallel (or antiparallel) to all nonzero columns of  $\hat{\Sigma}$ . This can only occur if all nonzero columns of  $\hat{\Sigma}$  are linearly dependent, which implies  $\text{Rank}[\hat{\Sigma}] \leq 1$ .

The Second Law of Drunk Dynamics can also be written in terms of infinitesimal covariance matrices. Let  $\boldsymbol{\nu}\boldsymbol{\nu}_m^T$  denote the outer product of  $\boldsymbol{\nu}_m$  with itself. Then (4.2) is:

$$\frac{d}{dt}|\bar{\mathbf{r}}|^2 = \sum_m \bar{\mathbf{r}}^T \left[ \boldsymbol{\nu}\boldsymbol{\nu}_m^T - |\boldsymbol{\nu}_m|^2 \hat{\mathbf{1}} \right] \bar{\mathbf{r}} = \bar{\mathbf{r}}^T \left( \hat{\Sigma}^T \hat{\Sigma} - \text{Tr}[\hat{\Sigma}^T \hat{\Sigma}] \hat{\mathbf{1}} \right) \bar{\mathbf{r}} = |\hat{\Sigma} \bar{\mathbf{r}}|^2 - |\bar{\mathbf{r}}|^2 \text{Tr}[\hat{\Sigma}^T \hat{\Sigma}]$$

If  $\bar{\mathbf{r}} = \mathbf{0}$  or  $\hat{\Sigma} = \hat{0}$ , then the radial velocity of  $\bar{\mathbf{r}}$  is zero. If neither  $\bar{\mathbf{r}}$  nor  $\hat{\Sigma}$  is zero, then let  $\{s_m\}$  denote the **singular values** of  $\hat{\Sigma}$ , i.e. the square roots of the eigenvalues of the positive-semidefinite matrix  $\hat{\Sigma}^T \hat{\Sigma}$ . The magnitude  $|\hat{\Sigma} \bar{\mathbf{r}}|$  is bounded from above by

$$|\hat{\Sigma} \bar{\mathbf{r}}|^2 \leq |\bar{\mathbf{r}}|^2 \max[s_m^2]$$

with equality only if  $\bar{\mathbf{r}}$  is an eigenvector of  $\hat{\Sigma}^T \hat{\Sigma}$  with eigenvalue  $\max[s_m^2]$ . The Hilbert-Schmidt inner product  $\text{Tr}[\hat{\Sigma}^T \hat{\Sigma}]$  is the squared **Frobenius norm** of  $\hat{\Sigma}$ :

$$\text{Frob}^2(\hat{\Sigma}) \equiv \text{Tr}[\hat{\Sigma}^T \hat{\Sigma}] = s_1^2 + s_2^2 + s_3^2$$

If  $\hat{\Sigma}$  has no more than one nonzero singular value, then  $\text{Frob}^2(\hat{\Sigma}) = \max[s_m^2]$ . Otherwise,  $\text{Frob}^2(\hat{\Sigma}) > \max[s_m^2]$ . The radial velocity of  $\bar{\mathbf{r}}$  is thus bounded from above by:

$$\frac{d}{dt}|\bar{\mathbf{r}}|^2 \leq |\bar{\mathbf{r}}|^2 \left( \max[s_m^2] - \text{Frob}^2[\hat{\Sigma}] \right) \leq 0$$

If  $\hat{\Sigma} \neq 0$  and  $\bar{\mathbf{r}} \neq \mathbf{0}$ , then  $\bar{\mathbf{r}}$  must drift inward unless  $\hat{\Sigma}$  has exactly one nonzero singular value  $s$  and  $\bar{\mathbf{r}}$  is an eigenvector of  $\hat{\Sigma}^T \hat{\Sigma}$  with eigenvalue  $s^2$ .

From a physical point of view, mean states gain von Neumann entropy until something “turns off most of the noise” and  $\text{Rank}[\hat{\Sigma}(t)]$  is reduced to 1 or 0. In these cases, the mean state can approach a nonzero steady state.<sup>6</sup> Otherwise  $\bar{\mathbf{r}}(t)$  continues towards the maximum-entropy mixture  $\mathbf{0}$  until we have no idea what state is actually present.

Equation (Master) fails for nonlinear models, so the results of this section are valid *only* for linear drunk models in which  $\boldsymbol{\mu}$  and  $\hat{\Sigma}$  both have no  $\mathbf{r}_t$  dependence. The surplus energy model in Section 4.6 was deliberately constructed as a nonlinear counterexample to the Second Law. Numerical simulations of this model appear to show the mean state approaching a finite-temperature thermal equilibrium state on the  $z$ -axis of the Bloch ball.

<sup>6</sup>The toy dephasing model in Section 4.3 is an example: it has a rank-1 volatility matrix, and the mean state drifts inward until it reaches the  $z$  axis and “gets stuck” at a nonzero steady-state value.

### 4.3 Example: Toy model of dephasing

This oversimplified model can be solved exactly, and the solution demonstrates many of the essential properties of drunk models, SF-It $\bar{o}$  conversion, and the Second Law. The solution here uses *It $\bar{o}$ 's Lemma*, the *Box Calculus*, and the *SF Chain Rule* from Appendix A. The model begins with a heuristic expression for a stochastic  $\mathbf{B}$  field in the  $z$  direction:

$$\mathbf{B}_t = (0, 0, \mu + \nu \frac{dW}{dt})$$

where  $\mu, \nu$  are real constants. The corresponding pre-equation is  $\frac{d}{dt}\mathbf{r}_t = -\mathbf{B}_t \times \mathbf{r}_t$ . Promote it to an SDE by “multiplying by  $dt$ ” and choosing Stratonovich-Fisk calculus:

$$\begin{aligned} d\mathbf{r}_t &= -(\mathbf{B}_t \times \mathbf{r}_t)dt = -(0, 0, \mu) \times \mathbf{r}_t dt - (0, 0, \nu) \times \mathbf{r}_t \circ dW \\ \begin{bmatrix} dx_t \\ dy_t \\ dz_t \end{bmatrix} &= \begin{bmatrix} 0 & \mu & 0 \\ -\mu & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} dt + \begin{bmatrix} 0 & \nu & 0 \\ -\nu & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} \circ dW \quad (\text{ToySF}) \end{aligned}$$

Geometric intuition suggests that  $\mathbf{r}_t$  rotates around the  $z$ -axis with a random angular frequency. Consider a rotation whose angle  $\phi_t$  solves the SDE  $d\phi_t = \mu dt + \nu \circ dW$ :

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \cos(\phi_t) & \sin(\phi_t) & 0 \\ -\sin(\phi_t) & \cos(\phi_t) & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \quad \phi_t \equiv \int_0^t \mu dt + \int_0^t \nu \circ dW = \mu t + \nu W_t$$

If this is correct, then  $\mathbf{r}_t$  rotates by a random angle  $\phi_t = \mu t + \nu W_t$  per time unit. To check the proposed solution, write  $x_t, y_t$  as functions of  $\phi_t$  and apply the SF Chain Rule:

$$\begin{aligned} dx(\phi_t) &= x'(\phi_t) \circ d\phi_t = [-x_0 \sin(\phi_t) + y_0 \cos(\phi_t)] \circ d\phi_t = y_t [\mu dt + \nu \circ dW] \\ dy(\phi_t) &= y'(\phi_t) \circ d\phi_t = [-x_0 \cos(\phi_t) - y_0 \sin(\phi_t)] \circ d\phi_t = -x_t [\mu dt + \nu \circ dW] \end{aligned}$$

This agrees with (ToySF), which verifies the proposed solution. The It $\bar{o}$  and SF solutions for  $\phi_t$  agree with each other, but the It $\bar{o}$  and SF solutions to (ToySF) do *not* agree. To see the distinction, apply It $\bar{o}$ 's Lemma to  $x(\phi_t)$  and  $y(\phi_t)$ :

$$\begin{aligned} dx(\phi_t) &= x'(\phi_t)d\phi_t + \frac{1}{2}x''(\phi_t)(d\phi_t)^2 \\ &= [-x_0 \sin(\phi_t) + y_0 \cos(\phi_t)]d\phi_t - \frac{1}{2}[x_0 \cos(\phi_t) + y_0 \sin(\phi_t)](d\phi_t)^2 \\ dy(\phi_t) &= y'(\phi_t)d\phi_t + \frac{1}{2}y''(\phi_t)(d\phi_t)^2 \\ &= [-x_0 \cos(\phi_t) - y_0 \sin(\phi_t)]d\phi_t + \frac{1}{2}[x_0 \sin(\phi_t) - y_0 \cos(\phi_t)](d\phi_t)^2 \end{aligned}$$

The Box Calculus rules  $dt^2 \rightarrow 0$ ,  $dt dW \rightarrow 0$ ,  $dW^2 \rightarrow dt$  imply  $(d\phi_t)^2 \rightarrow \nu^2 dt$ . (See Appendix A.) The functions  $x(\phi_t)$  and  $y(\phi_t)$  are solutions to the SF SDE (ToySF). According to Itô's Lemma, they are also solutions to the following Itô SDE:

$$\begin{aligned} dx_t &= y_t d\phi_t - \frac{1}{2}\nu^2 x_t dt = y_t(\mu dt + \nu dW) - \frac{1}{2}\nu^2 x_t dt \\ dy_t &= -x_t d\phi_t - \frac{1}{2}\nu^2 y_t dt = -x_t(\mu dt + \nu dW) - \frac{1}{2}\nu^2 y_t dt \end{aligned}$$

$$\begin{bmatrix} dx_t \\ dy_t \\ dz_t \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}\nu^2 & \mu & 0 \\ -\mu & -\frac{1}{2}\nu^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} dt + \begin{bmatrix} 0 & \nu & 0 \\ -\nu & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} dW \quad (\text{ToyIt}\bar{o})$$

The only differences between (ToySF) and (ToyIt $\bar{o}$ ) are the radial drift terms  $-\frac{1}{2}\nu^2 x_t dt$  and  $-\frac{1}{2}\nu^2 y_t dt$ . At first glance, these terms appear to pull  $\mathbf{r}_t$  inside the sphere, but  $|\mathbf{r}_t|$  is actually constant for all sample paths. For proof, define  $f(x, y, z) \equiv x^2 + y^2 + z^2$  and use Itô's Lemma to find  $df(x_t, y_t, z_t)$ :

$$\begin{aligned} df(\mathbf{r}) &= \nabla f(\mathbf{r}) \cdot d\mathbf{r} + \frac{1}{2} d\mathbf{r}^T [\hat{D}^2 f(\mathbf{r})] d\mathbf{r} \\ &= \begin{bmatrix} 2x_t & 2y_t \end{bmatrix} \begin{bmatrix} dx_t \\ dy_t \end{bmatrix} + \frac{1}{2} \begin{bmatrix} dx_t & dy_t \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} dx_t \\ dy_t \end{bmatrix} \\ &= 2x_t dx_t + 2y_t dy_t + (dx_t)^2 + (dy_t)^2 \end{aligned}$$

The Box Calculus deletes  $dt^2$  and  $dW dt$  terms, and the surviving terms cancel:

$$df(x_t, y_t, z_t) = (-\nu^2 x_t^2 - \nu^2 y_t^2 + \nu^2 x_t^2 + \nu^2 y_t^2) dt = 0$$

If the radial drift terms were removed from (ToyIt $\bar{o}$ ), then  $|\mathbf{r}_t|$  would grow exponentially and  $\mathbf{r}_t$  would spiral outward. This ‘‘centrifugal drift’’ is investigated in more detail in Appendix A. For now, the essential point is: the physical assumption  $|\mathbf{r}_t| = |\mathbf{r}_0|$  requires (ToySF) to be interpreted as an SF SDE. The Itô SDE (ToyIt $\bar{o}$ ) has the same solution, but it cannot be written in the heuristic form  $\frac{d}{dt}\mathbf{r}_t = -\mathbf{B}_t \times \mathbf{r}_t$ .

(Master) produces a master equation for the mean state coordinates  $\bar{x}(t), \bar{y}(t), \bar{z}(t)$ :

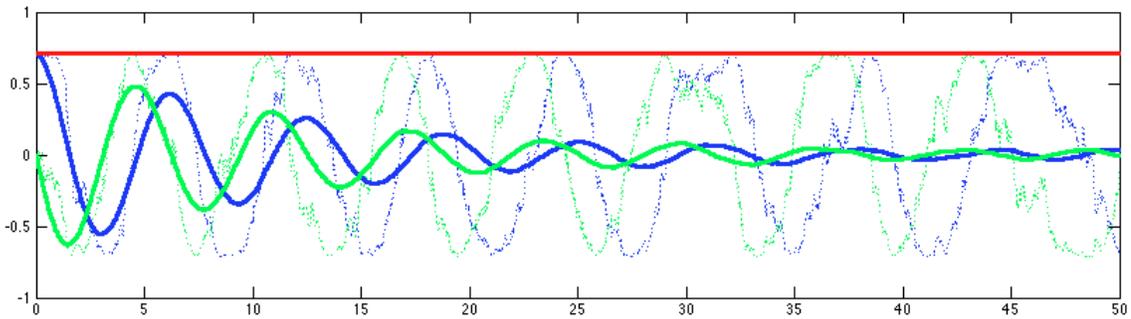
$$\begin{aligned} \frac{d}{dt}\bar{\mathbf{r}} &= -\mu \hat{K} \bar{\mathbf{r}} + \frac{1}{2}\nu^2 \hat{K}^2 \bar{\mathbf{r}} \\ \begin{bmatrix} d\bar{x}(t) \\ d\bar{y}(t) \\ d\bar{z}(t) \end{bmatrix} &= \begin{bmatrix} -\frac{1}{2}\nu^2 & \mu & 0 \\ -\mu & -\frac{1}{2}\nu^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \bar{x}(t) \\ \bar{y}(t) \\ \bar{z}(t) \end{bmatrix} dt \end{aligned}$$

Erasing the  $dW$  terms in (ToyItō) and replacing  $x_t, y_t, z_t$  with their expectation values  $\bar{x}(t), \bar{y}(t), \bar{z}(t)$  produces the same master equation. The  $z$  component  $\bar{z}(t)$  stays fixed. The other coordinates' expectation values are:

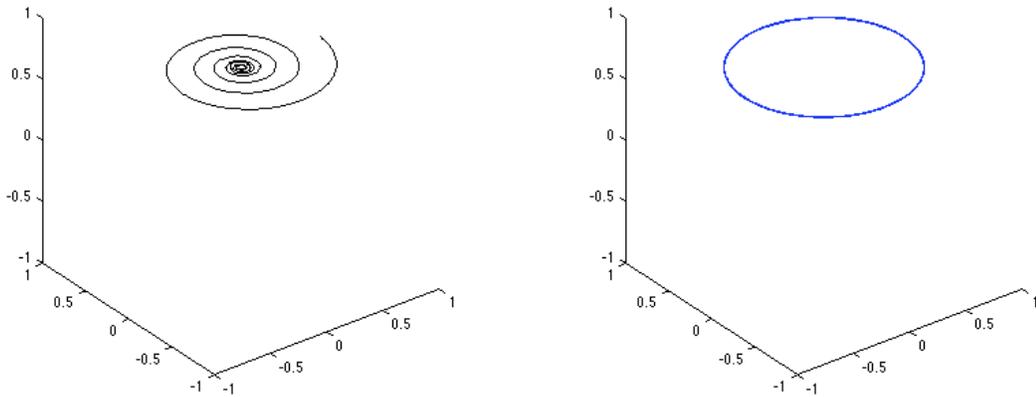
$$\begin{bmatrix} \bar{x}(t) \\ \bar{y}(t) \end{bmatrix} = e^{-\frac{1}{2}\nu^2 t} \begin{bmatrix} \cos(\mu t) & \sin(\mu t) \\ -\sin(\mu t) & \cos(\mu t) \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

Figure 4.1: Monte Carlo simulation of toy dephasing model.

Mean state is estimated by averaging 1000 simulations with  $\boldsymbol{\mu} = (0, 0, 1)$ ,  $\nu = 0.4$ ,  $t \in [0, 50]$ , and 50 timesteps per blink. Initial state is  $\mathbf{r}_0 = \frac{1}{\sqrt{2}}(1, 0, 1)$ .



Pauli coordinates  $x, y, z$  of mean state (solid lines) and a sample state (dotted lines).



Bloch-sphere trajectories of mean state (left) and a sample state (right).

Figure 4.1 shows numerical simulations of the toy dephasing model. On each trial, the true state  $\mathbf{r}_t$  rotates around the  $z$ -axis with an erratic angular velocity. After a short time  $\nu^2 t \ll 1$ ,  $\mathbf{r}_t$  is probably very close to where it should be. After a long time  $\nu^2 t \gg 1$ , the true angle  $\phi_t$  becomes a total mystery and the mean state approaches  $\bar{\mathbf{r}}(\infty) = (0, 0, z_0)$ .

The  $\bar{x}$  and  $\bar{y}$  coordinates of the mean state are the product of a sinusoidal oscillation and an exponential decay. This behavior is very similar to the *dephasing* which commonly occurs in qubit experiments. (Some examples are shown in Chapter 6.)

## 4.4 Example: Isotropic diffusion

In 1949, Yosida proved that stationary, isotropic Itô diffusion on a sphere is unique up to choice of time unit and initial condition.[60]<sup>7</sup> This process, sometimes called *spherical Brownian motion*, is another simplified model of qubit dissipation. The pre-equation assumes a constant mean field  $\boldsymbol{\mu} = \mathbf{0}$  and constant isotropic volatility  $\hat{\Sigma} = \nu \hat{1}$ :

$$\frac{d}{dt} \mathbf{r}_t = -\mathbf{B}_t \times \mathbf{r}_t = -\nu \left( \frac{dW_t^1}{dt}, \frac{dW_t^2}{dt}, \frac{dW_t^3}{dt} \right) \times \mathbf{r}_t$$

The superscripts  $dW_t^1, dW_t^2, dW_t^3$  are indices, not exponents. As before, “multiply both sides by  $dt$ ” and call the result an SF SDE:

$$\begin{bmatrix} dx_t \\ dy_t \\ dz_t \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \nu \\ 0 & -\nu & 0 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} \circ dW^1 + \begin{bmatrix} 0 & 0 & -\nu \\ 0 & 0 & 0 \\ \nu & 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} \circ dW^2 + \begin{bmatrix} 0 & \nu & 0 \\ -\nu & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} \circ dW^3$$

This equation can be written more concisely by using rotation generators  $\hat{I}, \hat{J}, \hat{K}$ .<sup>8</sup>

$$d\mathbf{r}_t = -\nu \left[ \hat{I} \circ dW^1 + \hat{J} \circ dW^2 + \hat{K} \circ dW^3 \right] \mathbf{r}_t \quad (\text{IsotropicSF})$$

The equivalent Itô SDE includes extra terms from the Wong-Zakai correction:

$$d\mathbf{r}_t = -\nu \left[ \hat{I} dW^1 + \hat{J} dW^2 + \hat{K} dW^3 \right] \mathbf{r}_t + \frac{1}{2} \nu^2 \left[ \hat{I}^2 + \hat{J}^2 + \hat{K}^2 \right] \mathbf{r}_t dt \quad (\text{IsotropicItô})$$

The matrix  $\hat{I}^2 + \hat{J}^2 + \hat{K}^2$  simplifies to  $-2\hat{1}$ . The master equation is given by (Master), or equivalently by erasing the  $dW$  terms in (IsotropicItô) and replacing  $\mathbf{r}_t$  with  $\bar{\mathbf{r}}(t)$ :

$$\frac{d}{dt} \bar{\mathbf{r}} = -\nu^2 \bar{\mathbf{r}} \quad \Rightarrow \quad \bar{\mathbf{r}}(t) = \bar{\mathbf{r}}(0) e^{-\nu^2 t} \quad (\text{IsotropicMaster})$$

Expectation values of all Pauli observables decay exponentially to  $\mathbf{r} = \mathbf{0}$ .

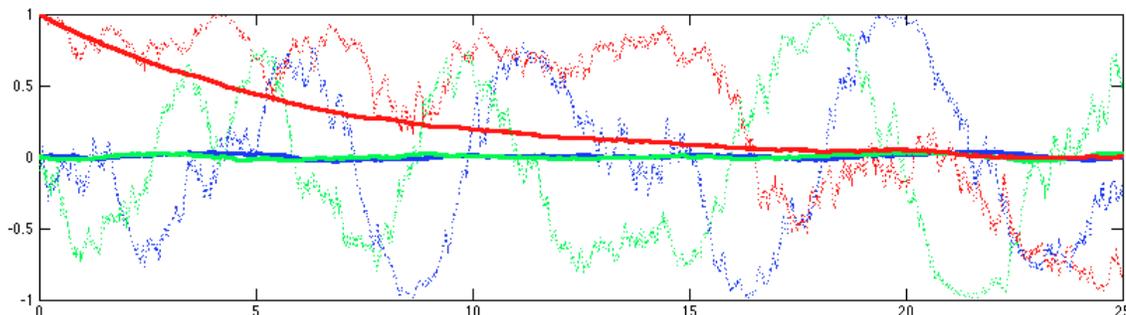
Appendix A finds the same result for  $\bar{\mathbf{r}}(t)$  by a much more laborious method which is closer to Yosida’s original work. The Fokker-Planck-Kolmogorov equation for the true state’s probability density  $p(t, \mathbf{r}) d^3 \mathbf{r}$  turns out to be the heat equation on a sphere. If the initial state is known to be some pure state  $\mathbf{r}_0$ , then  $p(0, \mathbf{r})$  has a Dirac  $\delta$  distribution. This is expanded as an infinite series of spherical harmonics, and the integral  $\bar{\mathbf{r}}(t) = E[\mathbf{r}_t] = \int \mathbf{r} p(t, \mathbf{r}) d^3 \mathbf{r}$  is calculated. When the dust settles, the result is  $\bar{\mathbf{r}}(t) = \mathbf{r}_0 e^{-\nu^2 t}$ .

<sup>7</sup>Yosida’s paper uses the term “3-sphere” to mean the surface  $|\mathbf{r}| = 1$  of a unit 3-ball  $|\mathbf{r}| \leq 1$ .

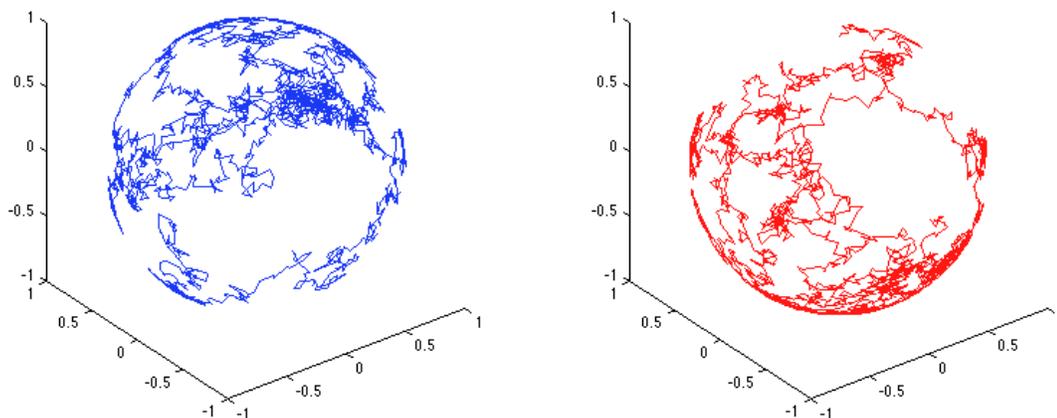
<sup>8</sup> $\hat{I}, \hat{J}, \hat{K}$  are *not* quaternions, though (IsotropicSF) and (IsotropicItô) do resemble quaternion equations.

Figure 4.2: Monte Carlo simulation of isotropic model.

Mean state is estimated by averaging 1000 simulations with  $\boldsymbol{\mu} = \mathbf{0}$ ,  $\nu = 0.4$ ,  $t \in [0, 25]$ , and 100 timesteps per blink. Initial state is  $\mathbf{r}_0 = (0, 0, 1)$ .



Pauli coordinates  $x$ ,  $y$ ,  $z$  of mean state (solid lines) and a sample state (dotted lines).



Bloch-sphere trajectories of two sample states.

Figure 4.2 shows numerical simulations of an isotropic model with the ground state as its initial condition. The isotropic model has no mean field; it is “all noise and no signal.” This model is not particularly realistic, but it can provide some useful geometric intuition.

On each trial, the true state wanders randomly around the Bloch sphere. After a short time  $\nu^2 t \ll 1$ , the true state is probably somewhere near its initial state. After a long time  $\nu^2 t \gg 1$ , every pure state is approximately equally probable and we have no idea what state is actually present. The trajectory of the mean state  $\bar{\mathbf{r}}$  is a straight line directly toward the maximum-entropy mixed state  $\mathbf{0}$ . (It is not shown in Figure 4.2.) The  $\bar{z}$  coordinate decays exponentially while the  $\bar{x}$  and  $\bar{y}$  coordinates remain approximately unchanged.

## 4.5 Example: Linear Bloch model

The **Bloch equation** is an equation of motion for nuclear magnetization.

$$\begin{aligned}\frac{d}{dt}m_x &= -\gamma(\mathbf{B} \times \mathbf{m})_x - \frac{m_x}{T_2} \\ \frac{d}{dt}m_y &= -\gamma(\mathbf{B} \times \mathbf{m})_y - \frac{m_y}{T_2} \\ \frac{d}{dt}m_z &= -\gamma(\mathbf{B} \times \mathbf{m})_z - \frac{m_z - m_\infty}{T_1}\end{aligned}$$

Here  $\mathbf{m}$  is the nuclear magnetization of some macroscopic chunk of material exposed to a magnetic field  $\mathbf{B}$ . (For notational clarity, time dependence of  $\mathbf{m}$  and  $\mathbf{B}$  is not shown.) The constant  $\gamma$  is the material's gyromagnetic ratio, and  $T_1, T_2$  are phenomenological constants called **relaxation times**. The constant  $m_\infty$  is the steady-state magnetization in the  $z$  direction. The notation  $(\mathbf{B} \times \mathbf{m})_x$  means the  $x$  component of  $\mathbf{B} \times \mathbf{m}$ , and similarly for  $y$  and  $z$ . In generator notation, the Bloch equation is:

$$\frac{d}{dt} \begin{bmatrix} m_x(t) \\ m_y(t) \\ m_z(t) \end{bmatrix} = \begin{bmatrix} -T_2^{-1} & \gamma B_z(t) & -\gamma B_y(t) \\ -\gamma B_z(t) & -T_2^{-1} & \gamma B_x(t) \\ \gamma B_y(t) & -\gamma B_x(t) & -T_1^{-1} \end{bmatrix} \begin{bmatrix} m_x(t) \\ m_y(t) \\ m_z(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ m_\infty/T_1 \end{bmatrix}$$

The time derivative of  $\mathbf{m}(t)$  is a linear function of  $\mathbf{m}(t)$  plus a constant affine term. The master equation for the model in this chapter has exactly the same form as the Bloch equation except for the affine term. The heuristic  $\mathbf{B}_t$  and pre-equation are:

$$\mathbf{B}_t = \boldsymbol{\mu}(t) + \hat{\Sigma} \frac{d\mathbf{W}}{dt} = \begin{bmatrix} \mu_x(t) \\ \mu_y(t) \\ \mu_z(t) \end{bmatrix} + \frac{d}{dt} \begin{bmatrix} \nu_1 & 0 & 0 \\ 0 & \nu_1 & 0 \\ 0 & 0 & \nu_z \end{bmatrix} \begin{bmatrix} dW^1 \\ dW^2 \\ dW^3 \end{bmatrix} \quad \frac{d}{dt} \mathbf{r}_t = -\mathbf{B}_t \times \mathbf{r}_t$$

The mean field  $\boldsymbol{\mu}$  is an arbitrary function of time.<sup>9</sup> The volatility matrix  $\hat{\Sigma}$  is diagonal and constant but not quite isotropic. As usual, “multiply the pre-equation by  $dt$ ” and consider the result to be an SF SDE. In terms of rotation generators  $\hat{I}, \hat{J}, \hat{K}$ , it is:

$$d\mathbf{r}_t = - \left[ \left( \mu_x(t) \hat{I} + \mu_y(t) \hat{J} + \mu_z(t) \hat{K} \right) dt + \nu_1 \hat{I} \circ dW^1 + \nu_1 \hat{J} \circ dW^2 + \nu_z \hat{K} \circ dW^3 \right] \mathbf{r}_t$$

<sup>9</sup> $\boldsymbol{\mu}$  is assumed to behave well enough that a unique solution to the ordinary LvN equation exists.

Using the (Master) formula, the resulting master equation is:

$$\frac{d}{dt}\bar{\mathbf{r}}(t) = \left[ -\mu_x(t)\hat{I} - \mu_y(t)\hat{J} - \mu_z(t)\hat{K} + \frac{1}{2}\left(\nu_1^2\hat{I}^2 + \nu_1^2\hat{J}^2 + \nu_z^2\hat{K}^2\right) \right] \bar{\mathbf{r}}(t)$$

In column-vector notation, the master equation is:

$$\frac{d}{dt} \begin{bmatrix} \bar{x}(t) \\ \bar{y}(t) \\ \bar{z}(t) \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}(\nu_1^2 + \nu_z^2) & \mu_z(t) & -\mu_y(t) \\ -\mu_z(t) & -\frac{1}{2}(\nu_1^2 + \nu_z^2) & \mu_x(t) \\ \mu_y(t) & -\mu_x(t) & -\nu_1^2 \end{bmatrix} \begin{bmatrix} \bar{x}(t) \\ \bar{y}(t) \\ \bar{z}(t) \end{bmatrix}$$

The following substitutions transform this into the Bloch equation (minus its affine term):

$$\mathbf{m}(t) \leftrightarrow \bar{\mathbf{r}}(t) \quad \gamma\mathbf{B}(t) \leftrightarrow \boldsymbol{\mu}(t) \quad T_1^{-1} \leftrightarrow \nu_1^2 \quad T_2^{-1} \leftrightarrow \frac{1}{2}(\nu_1^2 + \nu_z^2)$$

The toy dephasing model in section 4.3 predicts that mean states spiral inward exponentially with decay rate  $\frac{1}{2}\nu_z^2$ . With this in mind, define a **dephasing time**  $T_\phi^{-1} \equiv \frac{1}{2}\nu_z^2$ . This definition and the Bloch-equation substitutions lead to the following relation:

$$\frac{1}{T_2} = \frac{1}{2T_1} + \frac{1}{T_\phi}$$

The same equation has been found by superconducting qubit researchers using different methods.[47][61][62]  $T_1, T_2, T_\phi$  are variously called *relaxation*, *dephasing*, and/or *coherence* times by different authors. Positivity of  $\nu_1^2$  and  $\nu_z^2$  also implies  $T_2 \leq T_\phi$  and  $T_2 \leq 2T_1$ .

The linear Bloch model provides no instructions for determining  $\nu_1^2$  and  $\nu_z^2$  *a priori*. As Tanimura and Kubo cautioned, they are “merely a model appropriate for the problem rather than something to be derived.” Derivation of these parameters from physical considerations requires a more specific model of a qubit’s environment. However, experimental values of  $T_1$  and  $T_2$  can be used to define the **implied volatilities** of a qubit:

$$\tilde{\nu}_1 \equiv \frac{1}{\sqrt{T_1}} \quad \tilde{\nu}_z \equiv \sqrt{\frac{2}{T_2} - \tilde{\nu}_1^2} = \sqrt{\frac{2T_1 - T_2}{T_1T_2}}$$

Given empirical values for  $T_1, T_2$ , the implied volatilities  $\tilde{\nu}_1, \tilde{\nu}_z$  are what the parameters  $\nu_1$  and  $\nu_z$  “should be” if the linear Bloch model is accurate.<sup>10</sup>

<sup>10</sup>This idea is borrowed from a quantitative finance technique. Market prices of options are used to derive the volatility of the underlying asset, assuming the *Black-Scholes-Merton model* is correct. This implied volatility is then compared to empirical observations of the asset’s volatility.[63]

Figure 4.3: Monte Carlo simulation of Bloch model.

Mean state is estimated by averaging 1000 simulations with  $\boldsymbol{\mu} = (0, 0, 1)$ ,  $\nu_1 = 0.2$ ,  $\nu_z = 0.5$ ,  $t \in [0, 60]$ , and 50 timesteps per blink. Initial state is  $\mathbf{r}_0 = \frac{1}{\sqrt{2}}(1, 0, 1)$ .

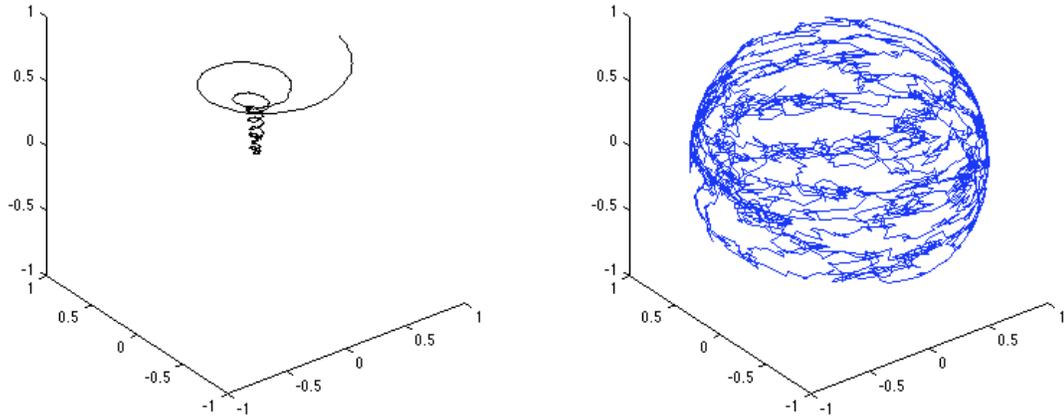
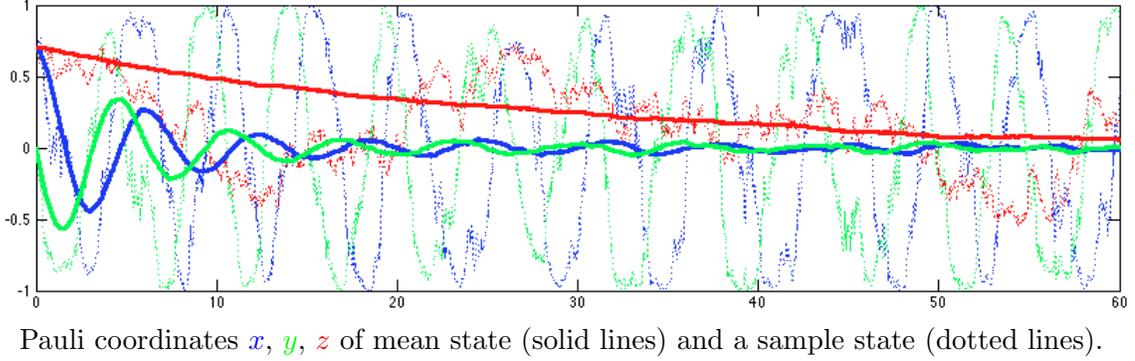


Figure 4.3 shows numerical simulations of a linear Bloch model qubit with mean field  $\boldsymbol{\mu} = (0, 0, 1)$  and initial condition  $(1, 0, 1)/\sqrt{2}$ . If there were no noise, the state would simply rotate about the  $z$ -axis. The sample state shown in Figure 4.3 is typical: it roughly rotates about the  $z$ -axis, but its  $z$ -coordinate “wobbles” and the rate of rotation is unsteady. Plotted as a 3D trajectory, the mean state spirals toward  $\mathbf{0}$ . Note that the mean  $\bar{z}$  coordinate decays more slowly than the mean  $\bar{x}$  and  $\bar{y}$  coordinates. In this example,  $T_1$  and  $T_2$  are:

$$T_1 = \frac{1}{\nu_1^2} = 25 \qquad T_2 = \frac{2}{\nu_1^2 + \nu_z^2} \approx 6.90$$

One benefit of the linear Bloch model model is its extreme generality. Every uncontrolled interaction between a generic qubit and the rest of the universe is represented by a few real numbers. (The toy dephasing model and isotropic model are special cases.) One weakness is the model’s inability to predict a thermal-equilibrium mean state as  $t \rightarrow \infty$ . The volatility matrix  $\hat{\Sigma}$  is not singular unless  $\nu_1$  or  $\nu_z$  is zero, so the Second Law of Drunk Dynamics does not allow a nonzero steady-state  $z$ -coordinate.

## 4.6 Example: Nonlinear surplus energy model

The nonlinear model in this section is not restricted by the Second Law. The drunken master equation is not valid for nonlinear models, but the numerical methods in Chapter 5 can still be used to run simulations. These simulations appear to show that a nonzero steady-state  $z$ -coordinate can indeed exist for nonlinear models.

From a physical point of view, linear models assume a qubit's Hamiltonian  $\mathbf{B}_t$  is statistically independent of its true state  $\mathbf{r}_t$ . Consider an analogous situation in which a swimmer uses a random model to predict the heights of waves she must swim through. If she is in the middle of an ocean, she might reasonably assume that the distribution of wave heights does not depend on her location. Linear drunk models assume an “ocean hypothesis” that uncontrolled environmental interactions do not depend on the qubit's state.

An extreme version of the analogy reveals another assumption. Suppose the swimmer's environment is a small bathtub. Waves now “remember” prior states in the sense that the swimmer's previous actions significantly affect the distribution of future wave heights. None of the noise models in this paper remember prior states. For nonlinear models,  $\mathbf{B}_t$  may depend on *where*  $\mathbf{r}_t$  is now, but it does not depend on *how*  $\mathbf{r}_t$  got there.<sup>11</sup>

The **surplus energy model** begins with the usual heuristic pre-equation:

$$\frac{d}{dt}\mathbf{r}_t = -\left(\boldsymbol{\mu} dt + \hat{\Sigma}(\mathbf{r}_t)\frac{d\mathbf{W}_t}{dt}\right) \times \mathbf{r}_t$$

For computational simplicity, the mean field  $\boldsymbol{\mu}$  is assumed constant. The volatility matrix  $\hat{\Sigma}(\mathbf{r}_t)$  does not explicitly depend on time, but it is a function of the true state  $\mathbf{r}_t$ :

$$\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \\ \mu \end{bmatrix} \quad \hat{\Sigma}(\mathbf{r}_t) = \begin{bmatrix} \nu_1 \mathcal{E}(\mathbf{r}_t) & 0 & 0 \\ 0 & \nu_1 \mathcal{E}(\mathbf{r}_t) & 0 \\ 0 & 0 & \nu_z \end{bmatrix} \quad \mathcal{E}(\mathbf{r}_t) \equiv \frac{z_\infty - z_t}{2|\mathbf{r}_t|}$$

Here  $\mu, \nu_1, \nu_z$ , and  $z_\infty \in [-1, 1]$  are real constants and  $\mathcal{E}(\mathbf{r}_t)$  is the **surplus energy function**. For linear models, the Second Law requires mean states to drift inward until something “turns off part of the noise” and  $\text{Rank}[\hat{\Sigma}] \leq 1$ . The surplus energy function does exactly that: whenever  $z_t = z_\infty$ , the volatility matrix has rank 1. This may appear to be an *ad hoc* assumption, and as far as the author knows, it is. It is presented here only to suggest that a nonzero steady-state solution may be possible for nonlinear drunk models.

<sup>11</sup>More precisely:  $\mathbf{B}_t$  is a Markovian process with respect to the natural filtration  $\mathcal{F}_t$  of  $\mathbf{r}_t$ .

The expected energy of a qubit with deterministic Hamiltonian  $\mathbf{B} = (0, 0, \mu)$  is:

$$\langle H \rangle = \text{Tr}[\hat{H}\hat{\rho}] = -\frac{1}{2}(\mathbf{B} \cdot \mathbf{r}) = -\frac{1}{2}\mu z_t$$

If  $z_\infty$  is the steady-state  $z$  coordinate, then the steady-state expected energy  $E_\infty$  is:

$$E_\infty = -\frac{1}{2}\mu z_\infty$$

The surplus energy function is  $(\langle H \rangle - E_\infty)$  divided by the qubit's energy gap:

$$\mathcal{E}(\mathbf{r}_t) \equiv \frac{\langle H \rangle - E_\infty}{E_1 - E_0} = \frac{-\frac{1}{2}\mu z_t + \frac{1}{2}\mu z_\infty}{\mu} = \frac{z_\infty - z_t}{2} = \frac{z_\infty - z_t}{2|\mathbf{r}_t|}$$

$\mathcal{E}(\mathbf{r}_t)$  can be negative, but this has no significance because multiplying a Wiener process by  $-1$  produces another Wiener process.[57][58]

The physical assumption that true states are pure requires  $|\mathbf{r}_t| = 1$  at all times, so the factor of  $|\mathbf{r}_t|$  is not strictly necessary. It is included in the definition only to ensure that  $\mathcal{E}(\mathbf{r}_t)$  can be easily written in spherical coordinates. If  $\theta_t$  and  $\theta_\infty$  denote the Bloch latitudes of states with  $z$ -coordinates  $z_t$  and  $z_\infty$ , then the surplus energy function is:

$$\mathcal{E}(\mathbf{r}_t) = \frac{z_\infty - z_t}{2|\mathbf{r}_t|} = \frac{1}{2} \left[ \cos(\theta_\infty) - \cos(\theta_t) \right]$$

From Section 2.5, the canonical density matrix for a qubit with energy gap  $\epsilon$  is:

$$\bar{\rho}_{\text{thermal}} = \frac{1}{2} \begin{bmatrix} 1 + \tanh(\beta\epsilon) & 0 \\ 0 & 1 - \tanh(\beta\epsilon) \end{bmatrix} \Leftrightarrow \bar{\mathbf{r}}_{\text{thermal}} = \left( 0, 0, \tanh(\beta\epsilon) \right) \quad \beta \equiv \frac{1}{kT}$$

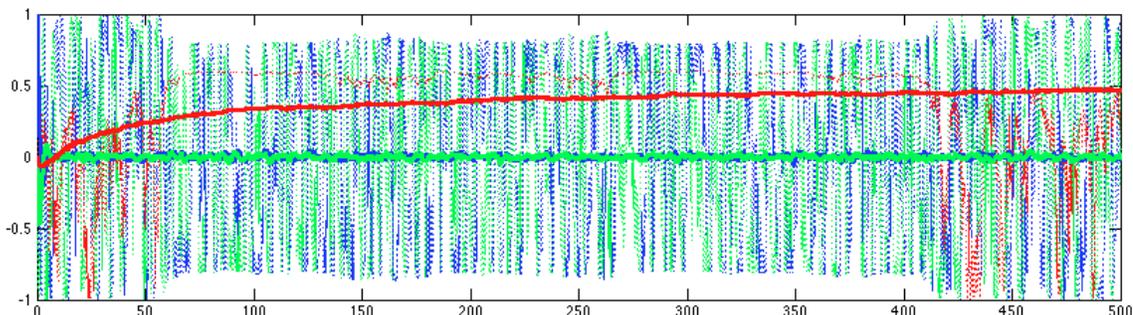
For a canonical mixed state, the equilibrium mean-state  $\bar{z}$  coordinate is  $z_\infty = \tanh(\beta\epsilon)$ .

The master equation from Section 4.1 cannot be trusted for nonlinear models, so finding the mean state may require solving an SDE. The ordinary deterministic LvN equation can be difficult to solve exactly if  $\mathbf{B}$  is not constant in time, and replacing  $\mathbf{B}$  with a stochastic process only increases the difficulty. Fortunately, the low dimensionality of the problem means Monte Carlo simulations can be performed without special hardware. (The numerical methods used are explained in detail in Chapter 5.)

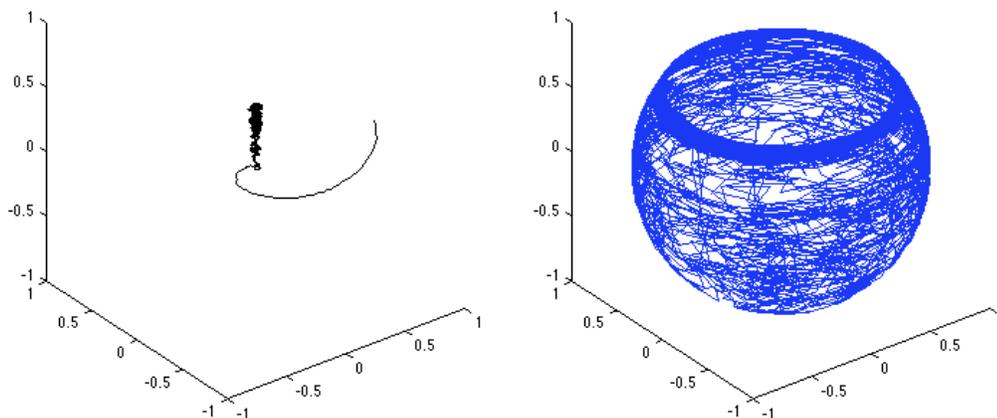
Figure 4.4 shows a simulation of the surplus energy model and the von Neumann entropy of the estimated mean state. For this example, the mean field is a constant  $\boldsymbol{\mu} = (0, 0, 1)$  and  $z_\infty = 0.6$ . At first, the mean state spirals inward toward  $\mathbf{0}$ . Flagrantly disregarding the Second Law, the entropy then slowly declines during a long “cooling off” period as the mean state moves upward along the  $z$ -axis toward  $(0, 0, z_\infty)$ .

Figure 4.4: Monte Carlo simulation of surplus energy model.

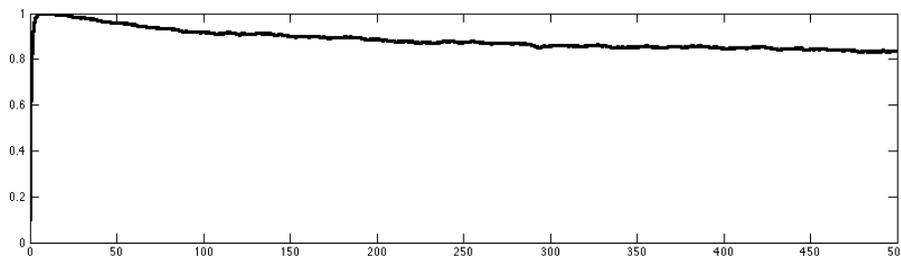
Mean state is estimated by averaging 1000 simulations with  $\boldsymbol{\mu} = (0, 0, 1)$ ,  $\nu_1 = \nu_z = 1$ ,  $t \in [0, 500]$ ,  $z_\infty = 0.6$ , and 20 timesteps per blink. Initial state is  $\mathbf{r}_0 = (1, 0, 0)$ .



Pauli coordinates  $x$ ,  $y$ ,  $z$  of mean state (solid lines) and a sample state (dotted lines).



Bloch-sphere trajectories of mean state (left) and a sample state (right).



Time series for Von Neumann entropy of mean state (bits).

The sample state in Figure 4.4 may give some hint of why the mean state slowly drifted north toward  $z_\infty$ . This state spent much of its time “stuck” near the latitude  $\theta_\infty$  where  $\mathcal{E}(\mathbf{r}_t) = 0$ . It did eventually break free, but it never entered the area of the Bloch sphere north of  $\theta_\infty$ . This behavior was typical of sample states for the surplus energy model. No states were seen crossing the equilibrium latitude  $\theta_\infty$ .

## 5 NUMERICAL METHODS

This chapter explains exponential numerical methods for ordinary and stochastic generator equations. In this thesis, an **ordinary generator equation** of dimension  $N$  is any first-order ordinary differential equation of the form

$$\dot{\mathbf{r}} = \hat{G}\mathbf{r}$$

where  $\mathbf{r}$  is an unknown  $1 \times N$  (real or complex) column-vector-valued function of time,  $\dot{\mathbf{r}}$  is its time derivative, and  $\hat{G}$  is a known  $N \times N$  (real or complex) matrix-valued function of  $\mathbf{r}$  and time. The name is borrowed from Lie group theory:  $\hat{G}$  is the *generator of time evolution* for the system in question. Generator equations are partitioned here into three categories, each of which is typically solved using different tactics.

- For **constant** generator equations,  $\hat{G}$  is constant.
- For **linear** generator equations,  $\hat{G}(t)$  may depend on time but not on  $\mathbf{r}$ .
- For **nonlinear** generator equations,  $\hat{G}(\mathbf{r}, t)$  depends on  $\mathbf{r}$  and possibly also  $t$ .

The Schrödinger equation for an  $N$ -level system is a complex linear generator equation.

$$\frac{d}{dt}|\Psi(t)\rangle = \hat{G}(t)|\Psi(t)\rangle \qquad \hat{G}(t) = -\frac{i}{\hbar}\hat{H}(t)$$

The LvN equation for Pauli coordinates of a qubit is a real linear generator equation.

$$\dot{\mathbf{r}}(t) = \hat{G}(t)\mathbf{r}(t) \qquad \hat{G}(t) = -[\mathbf{B}(t) \times] = -[B_x(t)\hat{I} + B_y(t)\hat{J} + B_z(t)\hat{K}]$$

Solutions to **linear** generator equations are linear transformations of an initial state:

$$\hat{G}(t) \text{ does not depend on } \mathbf{r} \quad \Rightarrow \quad \mathbf{r}(t) = \hat{T}(t, t_0)\mathbf{r}(t_0)$$

where the **time evolution operator**  $\hat{T}(t, t_0)$  is a matrix-valued function of  $t$  and  $t_0$ . If  $\hat{T}(t, t_0)$  can be found, then it can be re-used to find solutions with different initial conditions. (This is an example of the superposition principle for linear ODEs. For nonlinear generator equations, it may be necessary to run a new simulation for each initial condition.)

The density-matrix form of the Liouville-von Neumann equation can also be solved using generator-equation methods. Denote an initial density matrix by  $\bar{\rho}(t_0)$ . Construct an orthonormal basis  $\{|\psi_n\rangle\}$  of eigenvectors of  $\bar{\rho}(t_0)$ , and let  $\{\lambda_n\}$  denote the corresponding eigenvalues. Then  $\bar{\rho}(t_0)$  equals a convex combination of  $N$  pure states  $\{|\psi_n\rangle\langle\psi_n|\}$ :

$$\bar{\rho}(t_0) = \lambda_1|\psi_1\rangle\langle\psi_1| + \cdots + \lambda_N|\psi_N\rangle\langle\psi_N|$$

Find  $\hat{T}(t, t_0)$  for the Schrödinger equation. This operator sends each initial state  $|\psi_n\rangle$  to its new value  $\hat{T}(t, t_0)|\psi_n\rangle$ . Hiding the  $(t, t_0)$  dependence of  $\hat{T}$  for notational clarity,

$$\bar{\rho}(t) = \lambda_1\hat{T}|\psi_1\rangle\langle\psi_1|\hat{T}^\dagger + \cdots + \lambda_N\hat{T}|\psi_N\rangle\langle\psi_N|\hat{T}^\dagger = \hat{T}\bar{\rho}(t_0)\hat{T}^\dagger$$

For **constant** generator equations, the time evolution operator is a matrix exponential:

$$\hat{G} \text{ is constant} \quad \Rightarrow \quad \hat{T}(t, t_0) = \exp\left[(t - t_0)\hat{G}\right]$$

The exp symbol denotes the **matrix exponential function**:

$$\exp[\hat{M}] \equiv \sum_{l=0}^{\infty} \frac{1}{l!} \hat{M}^l = \hat{1} + \hat{M} + \frac{1}{2}\hat{M}^2 + \frac{1}{6}\hat{M}^3 + \cdots + \frac{1}{l!}\hat{M}^l + \cdots$$

This method can be generalized to **commutative** linear generator equations.

$$[\hat{G}(s), \hat{G}(t)] = 0 \text{ for all } s, t \quad \Rightarrow \quad \hat{T}(t, t_0) = \exp\left[\int_{t_0}^t \hat{G}(s) ds\right]$$

For noncommutative generator equations, approximations can be found using a *Dyson*, *Magnus*, or *Fer* series. Section 5.1 lists a few linear Magnus methods and the related *exponential midpoint method* (ExpMid) for nonlinear generator equations. These are especially useful for quantum simulations because they generate time evolution operators which are nearly unitary to very high precision.<sup>1</sup> Section 5.2 reviews matrix exponentials.

Section 5.3 briefly introduces *strong solvers* for stochastic differential equations. The strong solver used in this thesis combines the *Castell-Gaines strategy* in Subsection 5.3.1 with the ExpMid method and MATLAB's `expm` function for matrix exponentials. This method is relatively simple to implement and very good at keeping pure states pure. For example, all simulations in this thesis conserved Bloch-sphere radii to better than  $\pm 10^{-14}$ .

<sup>1</sup>Fer methods share this advantage but require more matrix exponentials per timestep.

## 5.1 Magnus methods

**Magnus methods** use the Magnus series to approximate solutions of generator equations. A time interval  $[t_0, t_J]$  is partitioned into  $J$  timesteps  $\{[t_j, t_j + h]\}$ , each of size  $h$ . (For simplicity, adaptive timestep-size algorithms are not used in this thesis.) The goal is then to find approximate short-term solutions in the form of a matrix exponential:

$$\mathbf{r}(t + h) \approx \exp[\hat{\Omega}(\mathbf{r}(t), t)]\mathbf{r}(t)$$

The first three terms of the Magnus series for  $\hat{\Omega}(t)$  are:[64]

$$\begin{aligned} \hat{\Omega}(t) = & \int_t^{t+h} \hat{G}(s) ds + \frac{1}{2} \int_t^{t+h} \int_t^s [\hat{G}(s), \hat{G}(r)] dr ds \\ & + \frac{1}{6} \int_t^{t+h} \int_t^s \int_t^r [\hat{G}(s), [\hat{G}(r), \hat{G}(q)]] + [\hat{G}(q), [\hat{G}(r), \hat{G}(s)]] dq dr ds + \dots \end{aligned}$$

Here  $q, r, s$  are dummy variables for time. For noncommutative generator equations, the series rapidly becomes a complicated mess.[65] Fortunately, it is possible to design an integrator with  $\mathcal{O}(h^6)$  global error using only 3 evaluations of  $\hat{G}(t)$   $m$  per timestep.[66]

A sufficient condition for the Magnus series to converge on a single timestep is:[67]

$$\int_{t_j}^{t_j+h} \|\hat{G}(s)\| ds < \pi$$

More precisely, “converges” means “converges in matrix norm” for some choice of matrix norm. For many practical purposes, the **Frobenius norm** is a reasonable choice.

$$\text{Frob}(\hat{M}) \equiv \sqrt{\text{Tr}[\hat{M}^\dagger \hat{M}]}$$

With this norm,  $\|\hat{M}\|^2$  is the sum of the absolute square of all matrix elements. The **spectral norm**, also known as the *induced Euclidean norm* or *largest singular value*, is arguably more relevant physically but also slower to calculate numerically.

$$\text{Spec}(\hat{M}) \equiv \sqrt{\lambda_{\max}(\hat{M}^\dagger \hat{M})}$$

Here  $\lambda_{\max}(\hat{M}^\dagger \hat{M})$  is magnitude of the largest-magnitude eigenvalue of  $\hat{M}^\dagger \hat{M}$ . If  $\hat{G}(t)$  has only imaginary eigenvalues and  $\|\hat{G}(t)\|$  is defined using the spectral norm, then sampling faster than the Nyquist rate of the solution is sufficient for convergence.<sup>2</sup>

<sup>2</sup>As noted in [67], series convergence is neither necessary nor sufficient for accurate simulations. Magnus methods sometimes produce accurate results even when the Nyquist condition is flagrantly violated. This “supersonic integration” is an active topic of research.[68][69]

Magnus methods are **geometric integrators** which preserve Lie group symmetries of the solution. In the context of the qubit LvN equation, these integrators keep states on the Bloch sphere. Specifically, linear combinations of rotation generators  $\hat{I}, \hat{J}, \hat{K}$  form a representation of the Lie algebra  $\mathfrak{so}(3)$ , which is closed under commutation:

$$[\hat{I}, \hat{J}] = \hat{K} \quad [\hat{J}, \hat{K}] = \hat{I} \quad [\hat{K}, \hat{I}] = \hat{J}$$

For the LvN equation,  $\hat{G}(t)$  is necessarily an element of  $\mathfrak{so}(3)$ . Every term in the Magnus series is also an element of  $\mathfrak{so}(3)$ . Suppose  $\Omega_l(t, t_0)$  is calculated by summing of the first  $l$  terms in the series, then ignoring the rest. The approximate time-evolution operator

$$\hat{T}_l(t, t_0) = \exp[\Omega_l(t, t_0)]$$

is then an element of  $SO(3)$ , the Lie group of 3D rotations. If the matrix exponential is calculated precisely enough, then even a low-accuracy Magnus method will evolve  $\mathbf{r}(t)$  by rotations and ensure that  $|\mathbf{r}(t)| = |\mathbf{r}(0)|$ . By a similar argument, Magnus integrators for state vectors of an  $N$ -level quantum system conserve total probability:[70]

$$\langle \Psi(t) | \Psi(t) \rangle = \langle \Psi(0) | \Psi(0) \rangle = 1$$

For many quantum simulations,  $|\Psi(t)\rangle$  follows a complicated path even if the behavior of the Hamiltonian  $\hat{H}(t)$  is relatively simple. (The Rabi flopping in Chapter 6 is a famous example.) Magnus methods often perform very well in this type of situation.[70][67]

These methods can also benefit greatly from parallelization. For linear systems, the calculations on each timestep do not require the results of previous timesteps. If  $J$  processors are available, then in principle all  $J$  timesteps can be calculated simultaneously.

A major weakness of Magnus methods is that matrix exponentials suffer from a “curse of dimensionality.” For a single qubit, the matrices involved are  $2 \times 2$  complex or  $3 \times 3$  real. For an  $Q$ -qubit system, a state vector has  $2^Q$  components. Simulating a 16-qubit system by directly applying the methods in this chapter to  $|\Psi\rangle$  would require exponentiating a  $65536 \times 65536$  complex matrix on each timestep. For a Monte Carlo simulation, the entire procedure would need to be repeated for each sample path. Even if symmetries are used to reduce the number of parameters, a very large number of numerical operations are typically needed for multi-qubit simulations. (As Feynman pointed out, any method for simulating high-dimensional quantum systems must overcome some version of this “curse.” [71])

### 5.1.1 ExpEuler and ExpMid methods

The  $\mathcal{O}(h)$  **exponential Euler method** approximates  $\hat{G}$  as constant on each step. If  $\mathbf{r}(t_j)$  is the state at time  $t_j$ , then the next state is calculated like so:

$$\mathbf{r}(t_j + h) \approx \exp \left[ h \hat{G}(\mathbf{r}(t_j), t_j) \right] \mathbf{r}(t_j)$$

For linear systems, the single-step time evolution operator  $\hat{S}_j$  is:

$$\hat{S}_j \equiv \exp \left[ h \hat{G}(t_j) \right] \quad \mathbf{r}(t_j + h) \approx \hat{S}_j \mathbf{r}(t_j)$$

These  $\{\hat{S}_j\}$  matrices can be multiplied to find an approximate time-evolution operator:

$$\hat{T}(t_J, t_0) = \hat{S}_{J-1} \hat{S}_{J-2} \cdots \hat{S}_1 \hat{S}_0 = \exp \left[ h \hat{G}(t_{J-1}) \right] \cdots \exp \left[ h \hat{G}(t_0) \right]$$

The linear ExpEuler method is a 1-term Magnus expansion with left-end quadrature:

$$\int_{t_j}^{t_j+h} \hat{G}(s) ds \approx h \hat{G}(t_j)$$

If all  $\hat{G}(t_j)$  matrices commute, then the exponential rule  $e^A e^B = e^{A+B}$  is valid. In the limit  $h \rightarrow 0$ ,  $J \rightarrow \infty$  with  $hJ = (t_J - t_0)$  fixed, ExpEuler becomes a (left) Riemann sum:

$$\hat{T}(t_J, t_0) = \lim_{h \rightarrow 0} \exp \left[ \sum_{j=0}^{J-1} h \hat{G}(t_0 + jh) \right] = \exp \left[ \int_{t_0}^{t_J} \hat{G}(s) ds \right]$$

The formula  $e^{A+B} = e^A e^B$  is *not* valid if  $AB \neq BA$ . For non-commuting matrices, it must be replaced with a *Baker-Campbell-Hausdorff* formula:[72][73][74][65]

$$\exp[\hat{A}] \cdot \exp[\hat{B}] = \exp \left[ \hat{A} + \hat{B} + \frac{1}{2}[\hat{A}, \hat{B}] + \frac{1}{12} \left( [\hat{A}, [\hat{A}, \hat{B}]] - [\hat{B}, [\hat{A}, \hat{B}]] \right) + \cdots \right]$$

where the dots indicate an infinite series of nested commutators.<sup>3</sup>

The  $\mathcal{O}(h^2)$  **exponential midpoint method** (ExpMid), also known as the **exponential Heun's method**, is more accurate than ExpEuler and only slightly more complicated.[76][77]<sup>4</sup> As the name suggests, it is an exponential version of the *midpoint method*. The generator

<sup>3</sup>For finite-dimensional quantum systems, a closed-form expression for the infinite series can be found.[75]

<sup>4</sup>Midpoint quadrature introduces error of at most  $\frac{1}{24}h^3 \|\partial_t^2 G(t_*)\|$ , where  $\|\partial_t^2 G(t_*)\|$  is the largest matrix spectral norm of  $\partial_t^2 \hat{G}(t)$  during the timestep. However, under certain mild assumptions about the commutators  $[\hat{G}(t), \hat{G}(s)]$ , Magnus methods tend to suppress quadrature errors.[67][70]

$\hat{G}$  is approximated as constant and sampled at the midpoint of each step:

$$\mathbf{r}(t_j + h) \approx \exp \left[ h \hat{G}(\mathbf{r}(t_j^\sharp), t_j^\sharp) \right] \mathbf{r}(t_j) \quad t_j^\sharp \equiv t_j + \frac{1}{2}h$$

Musicians may recognize the pun: the midpoint time  $t_j^\sharp$  is one half-step to the right of  $t_j$ . For linear generator equations, this method is no more costly than the Euler method:

$$\hat{S}_j \approx \exp \left[ h \hat{G}(t_j^\sharp) \right] \quad \hat{T}(t_J, t_0) \approx \hat{S}_{J-1} \hat{S}_{J-2} \cdots \hat{S}_1 \hat{S}_0$$

For nonlinear generator equations, the midpoint state  $\mathbf{r}(t_j^\sharp)$  is not known and must be estimated. ExpMid uses ExpEuler to estimate the midpoint state:

$$t_j^\sharp \equiv t_j + \frac{1}{2}h \quad \mathbf{r}_j^\sharp \equiv \exp \left[ \frac{1}{2}h \hat{G}(\mathbf{r}(t_j), t_j) \right] \mathbf{r}(t_j) \quad \mathbf{r}(t_j + h) \approx \exp \left[ h \hat{G}(\mathbf{r}_j^\sharp, t_j^\sharp) \right] \mathbf{r}(t_j)$$

The nonlinear ExpMid method requires *two* matrix exponentials per timestep rather than one. Both linear and nonlinear versions have global error  $\mathcal{O}(h^2)$ .

### 5.1.2 Advanced Magnus methods for linear ODEs

A few higher-order integrators for linear generator equations are below. Each method samples  $\hat{G}(t)$  at multiple **subsample times**, also called **collocation points**, per timestep.

---

$\mathcal{O}(h^4)$  with modified Gauss-Legendre quadrature:[76]

$$\begin{aligned} c_L &= \frac{1}{2} - \frac{\sqrt{3}}{6} & c_R &= \frac{1}{2} + \frac{\sqrt{3}}{6} \\ \hat{G}_L &= \hat{G}(t_j + c_L h) & \hat{G}_R &= \hat{G}(t_j + c_R h) \\ \Omega_j &= \frac{1}{2}h(\hat{G}_L + \hat{G}_R) - \frac{\sqrt{3}}{12}h^2[\hat{G}_L, \hat{G}_R] & \hat{S}_j &= \exp[\hat{\Omega}_j] \end{aligned}$$

---

$\mathcal{O}(h^4)$  with modified Simpson's Rule quadrature:[76]

$$\begin{aligned} \hat{G}_L &= \hat{G}(t_j) & \hat{G}_M &= \hat{G}(t_j + \frac{1}{2}h) & \hat{G}_R &= \hat{G}(t_j + h) \\ \hat{\Omega}_j &= \frac{1}{6}h(\hat{G}_L + 4\hat{G}_M + \hat{G}_R) - \frac{1}{12}h^2[\hat{G}_L, \hat{G}_R] & \hat{S}_j &= \exp[\hat{\Omega}_j] \end{aligned}$$

This method appears to calculate  $\hat{G}(t)$  three times per step rather than two. However, the right-end sample  $\hat{G}_R$  can be re-used on the next timestep. Only two function calls are needed on each timestep (except the first step). Because Simpson's Rule uses evenly-spaced subsample times, it may be easier to use with adaptive-timestep algorithms.

---

$\mathcal{O}(h^6)$  with modified Gauss-Legendre quadrature:[66]

$$\begin{aligned}
c_L &= \frac{1}{2} - \frac{\sqrt{15}}{10} & c_R &= \frac{1}{2} + \frac{\sqrt{15}}{10} \\
\hat{A}_L &= h\hat{G}(t_j + c_L h) & \hat{B}_1 &= h\hat{G}(t_j + \frac{1}{2}h) & \hat{A}_R &= h\hat{G}(t_j + c_R h) \\
\hat{B}_2 &= \frac{\sqrt{15}}{3}(\hat{A}_R - \hat{A}_L) & \hat{B}_3 &= \frac{10}{3}(\hat{A}_L - 2\hat{B}_1 + \hat{A}_R) \\
\hat{C}_1 &= [\hat{B}_1, \hat{B}_2] & \hat{D} &= 2\hat{B}_3 + \hat{C}_1 \\
\hat{C}_2 &= [\hat{B}_1, \hat{D}] & \hat{E} &= -20\hat{B}_1 - \hat{B}_3 + \hat{C}_1 & \hat{F} &= \hat{B}_2 - \frac{1}{60}\hat{C}_2 \\
\hat{C}_3 &= [\hat{E}, \hat{F}] & \hat{\Omega}_j &= \hat{B}_1 + \frac{1}{12}\hat{B}_3 + \frac{1}{240}\hat{C}_3 & \hat{S}_j &= \exp[\hat{\Omega}_j]
\end{aligned}$$

Higher-order Magnus methods for nonlinear generator equations are more complicated, require multiple matrix exponentials per timestep, and are not used in this thesis. For a detailed overview, see e.g. [77] and [76].

## 5.2 Matrix exponentials

The matrix exponential of an  $N \times N$  square matrix  $\hat{M}$  is:

$$\exp[\hat{M}] \equiv \sum_{l=0}^{\infty} \frac{1}{l!} \hat{M}^l = \hat{1} + \hat{M} + \frac{1}{2}\hat{M}^2 + \frac{1}{6}\hat{M}^3 + \dots + \frac{1}{l!}\hat{M}^l + \dots$$

Here  $\hat{1}$  indicates the  $N \times N$  identity matrix and  $\hat{M}^0 \equiv \hat{1}$  by definition. Some, but not all, properties of the function  $e^x$  remain valid for matrix exponentials:

$$(e^{\hat{M}})^{-1} = e^{-\hat{M}} \quad e^{(a+b)\hat{M}} = e^{a\hat{M}}e^{b\hat{M}} \quad \text{for all } a, b \in \mathbb{C}$$

The familiar rule  $e^{x+y} = e^x e^y$  does *not* work unless  $x$  and  $y$  commute:

$$e^{\hat{M}+\hat{N}} = e^{\hat{M}}e^{\hat{N}} \quad \text{if } [\hat{M}, \hat{N}] = 0$$

For matrices that commute, the following property also holds:

$$e^{\hat{N}}\hat{M}e^{-\hat{N}} = \hat{M} \quad \text{if } [\hat{M}, \hat{N}] = 0$$

Some additional properties not shared with the function  $e^x$  are:

$$(e^{\hat{M}})^\dagger = e^{\hat{M}^\dagger} \quad (e^{\hat{M}})^T = e^{\hat{M}^T} \quad \text{Det}[e^{\hat{M}}] = e^{\text{Tr}[\hat{M}]} \quad e^{\hat{N}}\hat{M}\hat{N}^{-1} = \hat{N}e^{\hat{M}}\hat{N}^{-1}$$

Every matrix commutes with its own exponential:

$$[\hat{M}, e^{a\hat{M}}] = [\hat{M}, \hat{1}] + a[\hat{M}, \hat{M}] + \frac{a^2}{2!}[\hat{M}, \hat{M}^2] + \frac{a^3}{3!}[\hat{M}, \hat{M}^3] + \dots = 0$$

If  $\hat{H}$  is self-adjoint and  $t$  is real, then  $\hat{U} = \exp[-t\hat{H}]$  is unitary:

$$\hat{U}^\dagger = \left(e^{-t\hat{H}}\right)^\dagger = e^{t\hat{H}^\dagger} = e^{t\hat{H}} = \left(e^{-t\hat{H}}\right)^{-1} = \hat{U}^{-1}$$

If  $\hat{A}$  is real antisymmetric and  $t$  is real, then  $\hat{R} = \exp[t\hat{A}]$  is orthogonal:

$$\hat{R}^T = \left(e^{t\hat{A}}\right)^T = e^{t\hat{A}^T} = e^{-t\hat{A}} = \hat{R}^{-1}$$

ExpMid and Magnus methods require a reliable method for numerical matrix exponentiation. This is a non-trivial task and has been the subject of several decades of research. Three of the “least dubious” methods, according to Moler and van Loan, are:[78]

1. Eigenvalue exponentiation
2. Taylor series with scaling and squaring
3. Padé approximants with scaling and squaring

**Eigenvalue exponentiation**, which requires a separate algorithm for matrix diagonalization, is reasonably fast and stable for normal matrices. A matrix  $\hat{M}$  is **normal** if it commutes with its adjoint:  $[\hat{M}, \hat{M}^\dagger] = 0$ . Normal matrices can be diagonalized by a unitary transformation:  $\hat{M} = \hat{U}\hat{D}\hat{U}^{-1}$ , where  $\hat{U}$  is unitary and  $\hat{D}$  is diagonal.

$$\hat{M} = \hat{U}\hat{D}\hat{U}^{-1} \quad \Rightarrow \quad e^{\hat{M}} = \hat{U}e^{\hat{D}}\hat{U}^{-1}$$

If  $\hat{U}$  is unitary, then  $\hat{U}^{-1} = \hat{U}^\dagger$  can be calculated quickly because  $\hat{U}^{-1} = \hat{U}^\dagger$ . Exponentiating  $\hat{D}$  is easy: exponentiate the diagonal elements and ignore the rest.

**Scaling and squaring** is a trick to ensure quick convergence and prevent catastrophic cancellation errors when summing a series. A natural number  $q$  is chosen such that  $\|\hat{M}\| < 2^q$ . The exponential of a rescaled matrix  $2^{-q}\hat{M}$  is then calculated and squared  $q$  times.

$$\exp[\hat{M}] = \left(\exp\left[\frac{1}{2^q}\hat{M}\right]\right)^{2^q}$$

The simulations in this thesis use MATLAB’s `expm` function, which uses Padé approximants with scaling and squaring. A truncated Taylor series with scaling and squaring is easier to code but typically slower by a factor between 1 and 2.[78] In testing, both of these methods were faster than diagonalization for matrices with spectral norm  $\|\hat{M}\| < 1$ .

### 5.3 Strong solvers for SDEs

A **strong solver** attempts to generate accurate approximations of sample paths of a stochastic differential equation. A **weak solver** attempts to accurately approximate the moments (mean, variance, skewness, kurtosis, etc.) of the solution.[79]

Plots of approximate sample paths can provide useful intuition about the geometric and qualitative properties of a system. They can also be used for Monte Carlo simulations of various properties of the system. For example, the estimated mean states plotted in Chapter 4 were found by simulating many sample paths and averaging them.

Unfortunately, simulating SDE sample paths is not as easy as simply inserting random numbers into an ODE solver. Consider a first-order Itô SDE of the form:

$$d\mathbf{r}_t = \mathbf{g}(\mathbf{r}_t, t) dt + \sum_{m=1}^M \boldsymbol{\eta}_m(\mathbf{r}_t, t) dW^m$$

The superscripts  $dW^m$  are indices, not exponents. The function  $\mathbf{g}$  is the **drift field** and the  $\{\boldsymbol{\eta}_m\}$  are **diffusion fields**. As before, break the time interval  $[t_0, t_J]$  into  $J$  equally-sized timesteps  $\{[t_j, t_j + h]\}$ . If the approximate strong solution depends only on samples of  $W_t$  taken at the sample times  $\{t_j\}$ , then its global error order is strictly limited. If the diffusion fields commute for all  $\mathbf{r}_t$  and  $t$ , then  $\mathcal{O}(h)$  is the best possible order; else it is  $\mathcal{O}(\sqrt{h})$ . [80] These error bounds cannot be avoided by adjusting  $\mathbf{g}$  to find an equivalent SF SDE.<sup>5</sup>

The  $\mathcal{O}(\sqrt{h})$  **Euler-Maruyama method** (EM) is one of the simplest strong solvers. Like the Euler method, it provides a useful introduction even though its accuracy and stability are unimpressive. In the small-timestep limit, EM converges to Itô solutions of an SDE. Let  $\Delta$  denote a (pseudo)random real number with a standard normal distribution. Then  $\sqrt{h}\Delta$  is a (pseudo)random real number with the same distribution as  $(W_{t+h} - W_t)$ . On each timestep, generate  $M$  independent samples  $\{\Delta_m\}$ . Replace each Wiener increment  $dW^m$  with  $\sqrt{h}\Delta_m$ . The new value of  $\mathbf{r}_{t+h}$  is approximated like so:

$$\mathbf{r}_{t+h} \approx \mathbf{r}_t + h \mathbf{g}(\mathbf{r}_t, t) + \sqrt{h} \sum_{m=1}^M \boldsymbol{\eta}_m(\mathbf{r}_t, t) \Delta_m$$

EM corresponds closely to Itô's definition of  $\int dW$ : it evaluates the drift and diffusion fields at the beginning of each timestep, just as Itô intended. For some applications, this is an advantage. But drunk model SDEs are typically easier to write in Stratonovich-Fisk form, which makes the Itô convergence of the EM method a nuisance. The EM method is also not a geometric integrator, and it is quite bad at keeping pure states on the Bloch sphere.

<sup>5</sup>Higher-order solvers require samples of iterated stochastic integrals, also called *Lévy areas*.

### 5.3.1 The Castell-Gaines strategy

The **Castell-Gaines strategy** (CG) is a class of strong SDE solvers introduced in [81]. The name *strategy* rather than *method* is used here because CG is can be implemented using any  $\mathcal{O}(h^2)$  ODE method. This is advantageous for drunk models because it can be used with exponential integrators which keep pure states pure.

CG is designed for Stratonovich-Fisk SDEs of the form:

$$d\mathbf{r}_t = \mathbf{g}(t, \mathbf{r}_t) dt + \sum_{m=1}^M \boldsymbol{\eta}_m(t, \mathbf{r}_t) \circ dW^m$$

On each timestep, the strategy is:

1. Define a temporary time coordinate  $s \in [0, 1]$  such that  $t = t_j + hs$ .
2. As with the EM method, generate  $M$  independent, standard-normally-distributed pseudorandom samples  $\{\Delta_m\}$ .
3. Construct the following *ordinary* differential equation:

$$\frac{d}{ds} \tilde{\mathbf{r}}(s) = h \mathbf{g}(\tilde{\mathbf{r}}(s), s) + \sqrt{h} \sum_{m=1}^M \boldsymbol{\eta}_m(\tilde{\mathbf{r}}(s), s) \Delta_m \quad (\text{CG ODE})$$

4. Set  $\tilde{\mathbf{r}}(0) = \mathbf{r}_{t_j}$ . Use an  $\mathcal{O}(h^2)$  ODE solver to approximate  $\tilde{\mathbf{r}}(1)$ . Set  $\mathbf{r}_{t_{j+1}} = \tilde{\mathbf{r}}(1)$ .

For noncommutative diffusion fields, Castell-Gaines methods have the same  $\mathcal{O}(\sqrt{h})$  strong global error as EM. However, they are generally more accurate than EM, and they converge to SF solutions instead of Itô solutions.

Castell and Gaines originally used Heun's method as their  $\mathcal{O}(h^2)$  ODE solver. The simulations in this thesis use ExpMid as the ODE solver. Calculating matrix exponentials costs extra time, but the return on this investment is that simulated states stay on the Bloch sphere to almost machine precision. For drunk qubits, (CG ODE) is:

$$\frac{d}{ds} \tilde{\mathbf{r}} = -h[\boldsymbol{\mu} \times] \tilde{\mathbf{r}} - \sqrt{h} \sum_{m=1}^3 \Delta_m [\boldsymbol{\nu}_m \times] \tilde{\mathbf{r}} \quad (\text{Drunk CG})$$

where  $\{\boldsymbol{\nu}_m\}$  are columns of the volatility matrix  $\hat{\Sigma}$ . (For notational clarity, any explicit  $s$  or  $\tilde{\mathbf{r}}(s)$  dependence is not shown.) On each timestep, generate a column  $\boldsymbol{\Delta}$  whose components are standard-normal pseudorandom samples  $\{\Delta_m\}$ . Then (Drunk CG) is:

$$\frac{d}{ds} \tilde{\mathbf{r}} = -\tilde{\mathbf{B}} \times \tilde{\mathbf{r}} = -(\tilde{B}_x \hat{I} + \tilde{B}_y \hat{J} + \tilde{B}_z \hat{K}) \tilde{\mathbf{r}} \quad \tilde{\mathbf{B}} \equiv h\boldsymbol{\mu} + \sqrt{h} \hat{\Sigma} \boldsymbol{\Delta}$$

For linear models, the linear ExpMid method can be used to solve (CG ODE):

$$t^\# \equiv t_j + \frac{1}{2}h \quad \tilde{\mathbf{B}}^\# \equiv h\boldsymbol{\mu}(t^\#) + \sqrt{h}\hat{\Sigma}(t^\#)\boldsymbol{\Delta} \quad \mathbf{r}_{t_j+h} \approx \exp \left[ -\tilde{B}_x^\# \hat{I} - \tilde{B}_y^\# \hat{J} - \tilde{B}_z^\# \hat{K} \right] \mathbf{r}_{t_j}$$

For nonlinear models,  $\tilde{\mathbf{B}}$  depends on  $\tilde{\mathbf{r}}$ , so its value at the midpoint time  $t^\#$  is unknown. The nonlinear ExpMid method uses ExpEuler to estimate  $\tilde{\mathbf{r}}$  and  $\tilde{\mathbf{B}}$  at the midpoint:

$$\begin{aligned} \tilde{\mathbf{B}} &\equiv h\boldsymbol{\mu}(\mathbf{r}_{t_j}, t_j) + \sqrt{h}\hat{\Sigma}(\mathbf{r}_{t_j}, t_j)\boldsymbol{\Delta} & \mathbf{r}^\# &\equiv \exp \left[ -\tilde{B}_x \hat{I} - \tilde{B}_y \hat{J} - \tilde{B}_z \hat{K} \right] \mathbf{r}_{t_j} \\ \tilde{\mathbf{B}}^\# &\equiv h\boldsymbol{\mu}(\mathbf{r}^\#, t^\#) + \sqrt{h}\hat{\Sigma}(\mathbf{r}^\#, t^\#)\boldsymbol{\Delta} & \mathbf{r}_{t_j+h} &\approx \exp \left[ -\tilde{B}_x^\# \hat{I} - \tilde{B}_y^\# \hat{J} - \tilde{B}_z^\# \hat{K} \right] \mathbf{r}_{t_j} \end{aligned}$$

The CG + ExpMid method corresponds fairly closely to the physical assumptions of a drunk model. On each timestep, the computer replaces each  $odW^m$  with a pseudorandom number  $\Delta_m$ . This simulated noise is added to the mean  $\mathbf{B}$  field, and the appropriate rotation on the Bloch sphere is calculated. Even if numerical errors are severe, each single-step time evolution operator is the exponential of a  $3 \times 3$  real antisymmetric matrix. Consequently, the new state  $\mathbf{r}_{t+h}$  is always an orthogonal transformation of the previous state  $\mathbf{r}_t$ .

## 6 SIMULATIONS VS. EXPERIMENTS

This chapter illustrates how drunk models can be used to simulate two experimental tests of qubit designs: *Rabi oscillations* and *Ramsey fringes*. These tests are used here as examples because similar experiments have been performed on a variety of qubit designs.

Section 6.1 is review and can be skipped or skimmed by readers familiar with NMR-type qubit controls. For ease of visualization and compatibility with modern experimental conventions, the relevant formulas are rewritten in terms of Pauli coordinates from Chapter 1. Sections 6.2 and 6.3 show numerical simulations of Rabi and Ramsey tests. Section 6.4 shows excerpts from four real qubit experiments for comparison.

Two types of simulations were performed using the MATLAB code in Appendix C:

- The master-equation method uses the `DrunkenMaster` script to numerically integrate the drunken master equation.
- The Monte Carlo method uses the `StochasticLinear` script to generate many pseudorandom possible states which are averaged to produce an approximate mean state.

These scripts are designed to be easily modified for different experiments and/or different drunk models. (Nonlinear models require the `StochasticNonlinear` script and cannot use the `DrunkenMaster` script.) The Monte Carlo method is more time-consuming because each simulation generates many sample paths of an SDE. The master-equation method only solves a single linear ODE, which is typically much faster and more accurate.

All simulations in this chapter use the linear Bloch model from Section 4.5 with various different volatilities. The master equation for this model is the Bloch equation with steady-state solution  $\mathbf{0}$  and decay times given by:

$$T_1 = \frac{1}{\nu_1^2} \quad T_2 = \frac{2}{\nu_1^2 + \nu_z^2} \quad T_\phi = \frac{2}{\nu_z^2}$$

$T_1$  is typically greater than  $T_\phi$  in practice, at least for superconducting qubits.[3] For this reason, most of the simulations in this chapter set  $\nu_z > \nu_1\sqrt{2}$ .

$$T_1 > T_\phi \quad \Leftrightarrow \quad \frac{1}{\nu_1^2} > \frac{2}{\nu_z^2} \quad \Leftrightarrow \quad 2\nu_1^2 < \nu_z^2$$

Rabi oscillations and Ramsey fringes are named for Rabi’s work on nuclear magnetic resonance and Ramsey’s related *separated oscillatory fields method*.<sup>[82][83]</sup> Rabi and Ramsey were awarded Nobel Prizes for their research in 1944 and 1989, respectively. Qubit tests based on Rabi cycles and Ramsey fringes later became an essential part of experiments such as <sup>[84]</sup> and <sup>[85]</sup> for which Haroche and Wineland were awarded the 2012 Nobel Prize.<sup>1</sup>

On each trial of a Rabi or Ramsey experiment, a qubit’s Hamiltonian is manipulated in a specific way, the qubit is allowed to evolve for some time  $t_j$ , and an energy measurement is performed. For each  $t_j$ , many trials are repeated and the number of excited-energy results is used to infer  $P(\text{excited}) = \frac{1}{2}(1 - z)$ . Quantum theory predicts approximately-sinusoidal oscillations of  $z$  and classical mechanics does not. In this way, Rabi and Ramsey experiments provide tests of the “quantumness” of a prototype qubit.

In real experiments, it is extremely common for the inferred  $z$ -coordinate to oscillate sinusoidally with an exponentially-decaying amplitude. Real qubits appear to decay smoothly from quantum to classical behavior with some characteristic decay times which are used to estimate  $T_1$  and  $T_2$  for that particular qubit. Section 6.4 shows several examples.

Similar damped-oscillator patterns of measurement probabilities are predicted by the drunken master equation. According to drunk models, the true state of a qubit never leaves the Bloch sphere, but noise causes the radius of the model’s mean state  $\bar{\mathbf{r}}(t)$  to decay exponentially. If a Rabi or Ramsey test is performed with many not-quite-identical trials, then the qubit’s inferred  $z$  coordinate will be close to the mean-state  $\bar{z}$  coordinate with high probability.

## 6.1 Pulse-controlled qubits

The theoretical description of a qubit in Chapter 1 is deliberately generic: a pure qubit state is a point on a sphere, and transformations between pure states are rotations. Real experiments require physical methods for producing these rotations. In practice, these methods are often more elaborate than simply choosing an axis and an angle.

Let  $\hat{H}_0$  denote the usual traceless reference Hamiltonian:

$$\hat{H}_0 = -\frac{1}{2}(E_1 - E_0)\hat{\sigma}_z = -\frac{\hbar}{2} \begin{bmatrix} \omega_0 & 0 \\ 0 & -\omega_0 \end{bmatrix}$$

As usual,  $\omega_0 = (E_1 - E_0)/\hbar$  is the qubit’s natural angular frequency. Suppose a qubit’s

---

<sup>1</sup>The cited experiments used the more complicated *Jaynes-Cummings model* in which a qubit is coupled to quantized electromagnetic field modes in a cavity.

Hamiltonian can be written as the sum of  $\hat{H}_0$  and an oscillating  $\hat{\sigma}_x$  term:

$$\hat{H}(t) = -\frac{1}{2}\hbar\omega_0\hat{\sigma}_z + A\hbar\cos(\omega_1t)\hat{\sigma}_x = -\frac{\hbar}{2} \begin{bmatrix} \omega_0 & 2A\cos(\omega_1t) \\ 2A\cos(\omega_1t) & -\omega_0 \end{bmatrix} \quad (\text{PCH})$$

In this thesis, this operator is called the **pulse control Hamiltonian** (PCH). In terms of Pauli coordinates and standardized spin- $\frac{1}{2}$  parameters, (PCH) is:

$$\mathbf{B}(t) = \left[ 2A\cos(\omega_1t), 0, \omega_0 \right]^T$$

The Schrödinger equation for a pulse-controlled qubit is:

$$\frac{d}{dt}|\Psi\rangle = -\frac{i}{\hbar}\hat{H}|\Psi\rangle \quad \Leftrightarrow \quad \begin{bmatrix} \dot{\alpha} \\ \dot{\beta} \end{bmatrix} = \frac{i}{2} \begin{bmatrix} \omega_0 & 2A\cos(\omega_1t) \\ 2A\cos(\omega_1t) & -\omega_0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

For clarity, time dependence of  $\hat{H}, |\Psi\rangle, \alpha, \beta$  is not shown. Unless otherwise noted, this chapter assumes  $\omega_1$  and  $\omega_0$  are positive. In many experiments,  $A$  is small and  $\omega_1$  is near resonance. More precisely, define the **detuning**  $\Delta \equiv \omega_1 - \omega_0$  and assume:

$$|\Delta| \ll |A| \ll |\omega_0|$$

If an experimenter has sufficiently precise control of  $A$  and  $\omega_1$ , then it is possible to rotate any initial pure state to a point arbitrarily close to any other pure state.

### 6.1.1 The rotating-wave approximation

In the **interaction picture**, new coordinates  $a, b$  are chosen by “rewinding” the time evolution generated by the reference Hamiltonian:

$$\hat{T}_0 = \exp\left[-\frac{it}{\hbar}\hat{H}_0\right] = \begin{bmatrix} e^{\frac{1}{2}i\omega_0t} & 0 \\ 0 & e^{-\frac{1}{2}i\omega_0t} \end{bmatrix} \quad \begin{bmatrix} a \\ b \end{bmatrix} \equiv \hat{T}_0^{-1} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \alpha e^{-\frac{1}{2}i\omega_0t} \\ \beta e^{\frac{1}{2}i\omega_0t} \end{bmatrix}$$

Time derivatives of these new interaction-picture coordinates  $a, b$  are:

$$\begin{bmatrix} \dot{a} \\ \dot{b} \end{bmatrix} = \begin{bmatrix} \dot{\alpha} e^{-\frac{1}{2}i\omega_0t} - \frac{1}{2}i\omega_0\alpha e^{-\frac{1}{2}i\omega_0t} \\ \dot{\beta} e^{\frac{1}{2}i\omega_0t} + \frac{1}{2}i\omega_0\beta e^{\frac{1}{2}i\omega_0t} \end{bmatrix}$$

Use the Schrödinger equation and (PCH) to find  $\dot{\alpha}$  and  $\dot{\beta}$ . Terms due to the reference Hamiltonian  $\hat{H}_0$  cancel, and what remains is:

$$\begin{bmatrix} \dot{a} \\ \dot{b} \end{bmatrix} = \imath A \begin{bmatrix} \beta \cos(\omega_1 t) e^{-\frac{1}{2}\imath\omega_0 t} \\ \alpha \cos(\omega_1 t) e^{\frac{1}{2}\imath\omega_0 t} \end{bmatrix} = \imath A \begin{bmatrix} b \cos(\omega_1 t) e^{-\imath\omega_0 t} \\ a \cos(\omega_1 t) e^{\imath\omega_0 t} \end{bmatrix} = \frac{\imath A}{2} \begin{bmatrix} b(e^{\imath(\omega_1 - \omega_0)t} + e^{\imath(-\omega_1 - \omega_0)t}) \\ a(e^{\imath(\omega_1 + \omega_0)t} + e^{\imath(-\omega_1 + \omega_0)t}) \end{bmatrix}$$

The **rotating-wave approximation** (RWA) discards the “fast”  $(\omega_1 + \omega_0)$  terms.[86]

$$\boxed{\begin{bmatrix} \dot{a} \\ \dot{b} \end{bmatrix} \approx \frac{\imath A}{2} \begin{bmatrix} b e^{\imath\Delta t} \\ a e^{-\imath\Delta t} \end{bmatrix}} \quad (\text{RWA})$$

This 2-dimensional 1st-order ODE can be rewritten as a 1-dimensional 2nd-order ODE which can be solved exactly. Take the time derivative of  $\dot{b}$  to find:

$$\ddot{b} = \frac{1}{2}\imath A (a e^{-\imath\Delta t} - \imath\Delta a e^{\imath\Delta t}) = -\frac{1}{4}A^2 b - \imath\Delta \dot{b}$$

This is a linear homogenous 2nd-order ODE, so look for solutions of the form  $e^{\lambda t}$ .

$$\lambda^2 = -\frac{1}{4}A^2 - \imath\Delta\lambda \quad \lambda = \frac{1}{2} \left( -\imath\Delta \pm \sqrt{-\Delta^2 - A^2} \right)$$

Define the **Rabi flopping frequency**  $\Omega$ :

$$\boxed{\Omega \equiv \sqrt{\Delta^2 + A^2} = \sqrt{(\omega_1 - \omega_0)^2 + A^2}}$$

The general solution for  $b(t)$  is a linear combination of two oscillating modes:

$$b(t) = C_1 e^{-\frac{1}{2}\imath(\Delta + \Omega)t} + C_2 e^{-\frac{1}{2}\imath(\Delta - \Omega)t}$$

The constants  $C_1, C_2$  are determined by initial conditions. The excitation probability  $\|b(t)\|^2$  “flops” up and down at the Rabi flopping frequency  $\Omega$ :

$$\|b(t)\|^2 = \|C_1\|^2 + \|C_2\|^2 + C_1^* C_2 e^{\imath\Omega t} + C_2^* C_1 e^{-\imath\Omega t}$$

If the qubit is initially in its ground state, then  $a(0) = 1$  (neglecting overall phase factors) and  $b(0) = 0$ . These initial conditions require  $b(0) = C_1 + C_2 = 0$ . The initial  $\dot{b}(0)$  is:

$$\dot{b}(0) = -\frac{1}{2}\imath [(\Delta + \Omega)C_1 + (\Delta - \Omega)C_2] = -\frac{1}{2}\imath [\Delta(C_1 + C_2) + \Omega(C_1 - C_2)] = \frac{1}{2}\imath\Omega(C_2 - C_1)$$

$$\dot{b}(0) = \frac{1}{2}iAa(0) = \frac{1}{2}iA \quad \Leftrightarrow \quad \frac{1}{2}i\Omega(C_2 - C_1) = \frac{1}{2}iA \quad \Leftrightarrow \quad C_2 = \frac{A}{2\Omega} = -C_1$$

The exact solution for  $b(t)$  with initial conditions  $a(0) = 1, b(0) = 0$  is:

$$b(t) = \frac{A}{2\Omega} \left[ -e^{-\frac{1}{2}i(\Delta+\Omega)t} + e^{-\frac{1}{2}i(\Delta-\Omega)t} \right] = \frac{iA}{\Omega} e^{-\frac{1}{2}i\Delta t} \sin\left(\frac{1}{2}\Omega t\right)$$

The probability of detecting the excited energy  $E_1$  at time  $t$  is:

$$\|\beta(t)\|^2 = \|b(t)\|^2 = \frac{A^2}{\Omega^2} \sin^2\left(\frac{1}{2}\Omega t\right) = \frac{A^2}{2\Omega^2} [1 - \cos(\Omega t)]$$

In the limit of perfect resonance,  $\omega_1 = \omega_0$ . In this case,  $\Delta = 0$ ,  $\Omega = |A|$ , and

$$\|\beta(t)\|^2 = \frac{1}{2} [1 - \cos(\Omega t)] \quad \Leftrightarrow \quad z(t) = \cos(\Omega t)$$

In this way, a resonant pulse can be used to set the excitation probability of a qubit. For example, a pulse of duration  $t_\pi \equiv \pi/\Omega$  sends the ground state to the excited state. Applying the pulse for twice as long produces one full **Rabi cycle**  $|0\rangle \rightarrow |1\rangle \rightarrow |0\rangle$ .

### 6.1.2 A geometric view: $\pi$ pulses and spiral operators

Pauli coordinates provide some geometric intuition about the RWA. The reference Hamiltonian is  $\mathbf{B}_0 = (0, 0, \omega_0)$  and the LvN equation is  $\dot{\mathbf{r}} = -\mathbf{B} \times \mathbf{r}$ . When no pulse is applied, the qubit's state  $\mathbf{r}$  rotates about the  $z$  axis with angular velocity  $-\omega_0$ .

$$\mathbf{r}(t) = \hat{T}_0(t)\mathbf{r}(0) \quad \Leftrightarrow \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos(\omega_0 t) & \sin(\omega_0 t) & 0 \\ -\sin(\omega_0 t) & \cos(\omega_0 t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$$

Define interaction-picture coordinates  $\mathbf{q} = (q_x, q_y, q_z)$  by “rewinding” this rotation:

$$\mathbf{q}(t) \equiv \hat{T}_0^{-1}(t)\mathbf{r}(t)$$

$$\begin{bmatrix} q_x \\ q_y \\ q_z \end{bmatrix} \equiv \begin{bmatrix} \cos(\omega_0 t) & -\sin(\omega_0 t) & 0 \\ \sin(\omega_0 t) & \cos(\omega_0 t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \cos(\omega_0 t) - y \sin(\omega_0 t) \\ x \sin(\omega_0 t) + y \cos(\omega_0 t) \\ z \end{bmatrix}$$

This  $\mathbf{q}$  is what the qubit's Bloch vector would look like from the point of view of an observer co-rotating with the qubit's natural angular frequency. Note that  $q_z$  is just the original  $z$

coordinate. The time derivatives of the other two  $q$  coordinates are:

$$\begin{bmatrix} \dot{q}_x \\ \dot{q}_y \end{bmatrix} = \begin{bmatrix} \dot{x} \cos(\omega_0 t) - \omega_0 x \sin(\omega_0 t) - \dot{y} \sin(\omega_0 t) - \omega_0 y \cos(\omega_0 t) \\ \dot{x} \sin(\omega_0 t) + \omega_0 x \cos(\omega_0 t) + \dot{y} \cos(\omega_0 t) - \omega_0 y \sin(\omega_0 t) \end{bmatrix}$$

When a pulse is applied, the Hamiltonian is (PCH):

$$\mathbf{B}(t) = \left[ 2A \cos(\omega_1 t), 0, \omega_0 \right]^T$$

Time evolution is given by the LvN equation  $\dot{\mathbf{r}} = -\mathbf{B} \times \mathbf{r}$ :

$$\dot{\mathbf{r}} = -\left[ 2A \cos(\omega_1 t) \hat{I} + \omega_0 \hat{K} \right] \mathbf{r}$$

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} 0 & \omega_0 & 0 \\ -\omega_0 & 0 & 2A \cos(\omega_1 t) \\ 0 & -2A \cos(\omega_1 t) & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Substituting  $\dot{x}, \dot{y}, \dot{z}$  into the equation for  $\dot{\mathbf{q}}$  and cancelling terms,

$$\begin{bmatrix} \dot{q}_x \\ \dot{q}_y \\ \dot{q}_z \end{bmatrix} = 2A \begin{bmatrix} -z \cos(\omega_1 t) \sin(\omega_0 t) \\ z \cos(\omega_1 t) \cos(\omega_0 t) \\ -y \cos(\omega_1 t) \end{bmatrix} = 2A \begin{bmatrix} -q_z \cos(\omega_1 t) \sin(\omega_0 t) \\ q_z \cos(\omega_1 t) \cos(\omega_0 t) \\ q_x \cos(\omega_1 t) \sin(\omega_0 t) - q_y \cos(\omega_1 t) \cos(\omega_0 t) \end{bmatrix}$$

To apply the RWA, ignore any “fast” ( $\omega_0 + \omega_1$ ) terms:

$$\begin{aligned} \cos(\omega_1 t) \sin(\omega_0 t) &= \frac{1}{4i} \left( e^{i(\omega_0 + \omega_1)t} - e^{-i(\omega_0 + \omega_1)t} - e^{i\Delta t} + e^{-i\Delta t} \right) \approx -\frac{1}{2} \sin(\Delta t) \\ \cos(\omega_1 t) \cos(\omega_0 t) &= \frac{1}{4} \left( e^{i(\omega_0 + \omega_1)t} + e^{-i(\omega_0 + \omega_1)t} + e^{i\Delta t} + e^{-i\Delta t} \right) \approx \frac{1}{2} \cos(\Delta t) \end{aligned}$$

The RWA in interaction-picture  $\mathbf{q}$  coordinates is then:

$$\boxed{\begin{bmatrix} \dot{q}_x \\ \dot{q}_y \\ \dot{q}_z \end{bmatrix} \approx A \begin{bmatrix} 0 & 0 & \sin(\Delta t) \\ 0 & 0 & \cos(\Delta t) \\ -\sin(\Delta t) & -\cos(\Delta t) & 0 \end{bmatrix} \begin{bmatrix} q_x \\ q_y \\ q_z \end{bmatrix}} \quad (\text{RWA})$$

In terms of rotation generators, (RWA) can be written more compactly:

$$\dot{\mathbf{q}} = -A \left[ \cos(\Delta t) \hat{I} - \sin(\Delta t) \hat{J} \right] \mathbf{q}$$

If  $|\Delta| \ll |A|$ , then the short-term behavior of  $\mathbf{q}$  is rotation about the  $x$  axis with angular velocity  $\approx -A$ . On longer timescales, the axis of rotation falls out of sync with the  $q_x$  axis and slowly rotates about the equator.

In the limit of perfect resonance  $\Delta \rightarrow 0$ , (RWA) is a constant generator equation. Its exact solution is a matrix exponential which rotates  $\mathbf{q}$  about the  $q_x$  axis:

$$\mathbf{q}(t) = \hat{R}(t)\mathbf{q}(0) \quad \hat{R}(t) = \exp[-At\hat{I}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(At) & \sin(At) \\ 0 & -\sin(At) & \cos(At) \end{bmatrix}$$

From a non-rotating point of view, the  $q_x$  axis itself is rotating around the equator with angular velocity  $\omega_0$ . To calculate  $\mathbf{r}(t)$ , transform back to non-rotating coordinates:

$$\mathbf{r}(t) = \hat{T}(t)\mathbf{q}(t) = \hat{T}(t)\hat{R}(t)\mathbf{q}(0) = \hat{T}(t)\hat{R}(t)\mathbf{r}(0)$$

Time evolution of  $\mathbf{r}(t)$  is given by the **spiral operator**  $\hat{S}(t) \equiv \hat{T}(t)\hat{R}(t)$ .

$$\hat{S}(t) = \begin{bmatrix} \cos(\omega_0 t) & \sin(\omega_0 t) \cos(At) & \sin(\omega_0 t) \sin(At) \\ -\sin(\omega_0 t) & \cos(\omega_0 t) \cos(At) & \cos(\omega_0 t) \sin(At) \\ 0 & -\sin(At) & \cos(At) \end{bmatrix}$$

If a weak resonant pulse is applied for a time  $t$  to a qubit state  $\mathbf{r}_0$ , then the spiral operator calculates the new state  $\mathbf{r}(t) = \hat{S}(t)\mathbf{r}_0$  according to the RWA.

A **pi pulse** is a pulse with duration  $\pi/|A|$ , weak amplitude  $|A| \ll \omega_0$ , and negligible detuning  $|\omega_1 - \omega_0| \ll |A|$ . The spiral operator for a  $\pi$  pulse is:

$$\hat{S}_\pi = \begin{bmatrix} \cos(\omega_0 t_\pi) & -\sin(\omega_0 t_\pi) & 0 \\ -\sin(\omega_0 t_\pi) & -\cos(\omega_0 t_\pi) & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad t_\pi \equiv \frac{\pi}{|A|}$$

This operator sends a state at the North pole  $(0, 0, 1)$  to the South pole  $(0, 0, -1)$ . Figure 6.1 shows an example  $\pi$  pulse and an illustration from Hahn's 1953 paper on the *spin echo* technique.[87] For this example, the initial state is  $\mathbf{r}_0 = (0, 0, 1)$ . The pulse-control Hamiltonian is used with  $A = \frac{1}{20}$ ,  $\omega_1 = \omega_0 = 1$ . The simulation is run for  $t \in [0, 20\pi]$ .

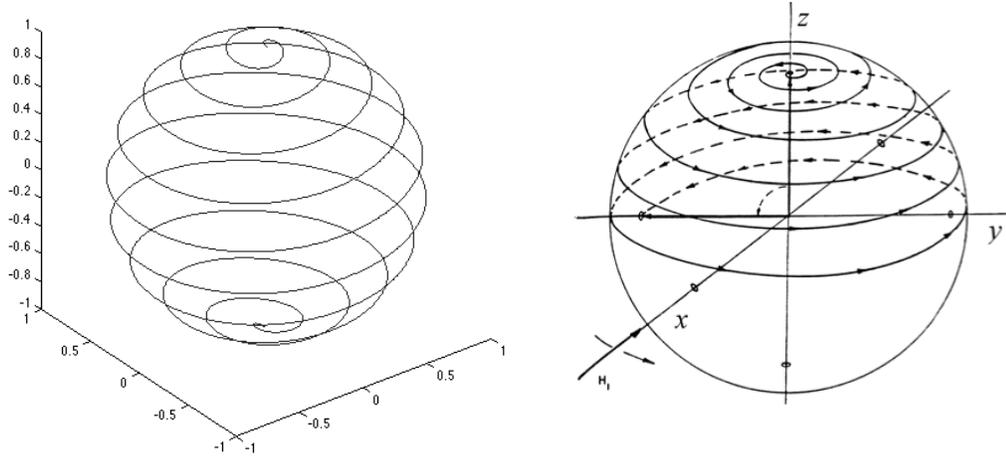


Figure 6.1: Simulated  $\pi$  pulse and illustration of a  $\pi/2$  pulse.

Left: Results of `DrunkenMaster` with zero volatility, pulse amplitude  $|A| = \frac{1}{20}$ , and perfectly-resonant pulse frequency  $\omega_1 = \omega_0 = 1$ . Right: Hahn's illustration from [87].

It is sometimes convenient to also define  $\pi/2$  pulses and  $2\pi$  pulses with durations:

$$t_{\frac{\pi}{2}} \equiv \frac{\pi}{2|A|} \quad t_{\pi} \equiv \frac{\pi}{|A|} \quad t_{2\pi} \equiv \frac{2\pi}{|A|}$$

For pulses whose duration is not an integer multiple of  $t_{\pi}$ , the spiral operator may depend on the sign of  $A$ . For a  $\pi/2$  pulse, it is:

$$\hat{S}_{\frac{\pi}{2}} = \begin{bmatrix} \cos(\frac{1}{2}\omega_0 t_{\pi}) & 0 & \pm \sin(\frac{1}{2}\omega_0 t_{\pi}) \\ -\sin(\frac{1}{2}\omega_0 t_{\pi}) & 0 & \pm \cos(\frac{1}{2}\omega_0 t_{\pi}) \\ 0 & \mp 1 & 0 \end{bmatrix} \quad \pm = \begin{cases} + & \text{if } A > 0 \\ - & \text{if } A < 0 \end{cases}$$

Figure 6.2 shows one full Rabi cycle (a  $2\pi$  pulse) with  $A = \frac{1}{20}$  and  $\omega_1 = \omega_0 = 1$ . Small wobbles in the  $z$  coordinate are caused by the “fast” terms neglected in the RWA.

Figure 6.3 shows detuned Rabi cycles with  $\omega_1 = 0.999\omega_0$ . Interaction-picture coordinates  $\mathbf{q}$  are shown in a rotating frame. Schrödinger-picture coordinates  $\mathbf{r}$  are shown in a non-rotating frame. The RWA was used to calculate  $\mathbf{q}$  but not  $\mathbf{r}$  in this figure.

Figures 6.1, 6.2, and 6.3 were made by the `DrunkenMaster` script, which does *not* rely on the RWA. To simulate an ideal pulse-controlled qubit with `DrunkenMaster`, set the volatility matrix `Sigma` to zero and the mean field `Mu` to  $\boldsymbol{\mu}(t) = \mathbf{B}(t) = [2A \cos(\omega_1 t), 0, \omega_0]^T$ .

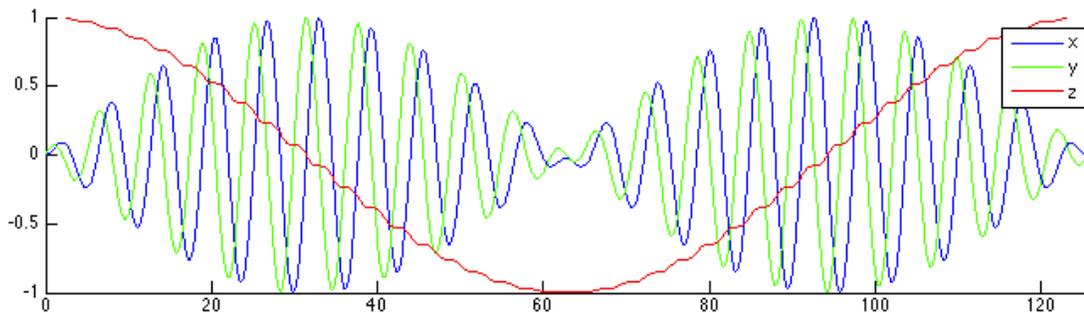


Figure 6.2: Full Rabi cycle with  $A = \frac{1}{20}$ ,  $\omega_1 = \omega_0 = 1$ , and  $t \in [0, 40\pi]$ . Small “wobbles” in the  $z$ -coordinate are neglected by the rotating-wave approximation.

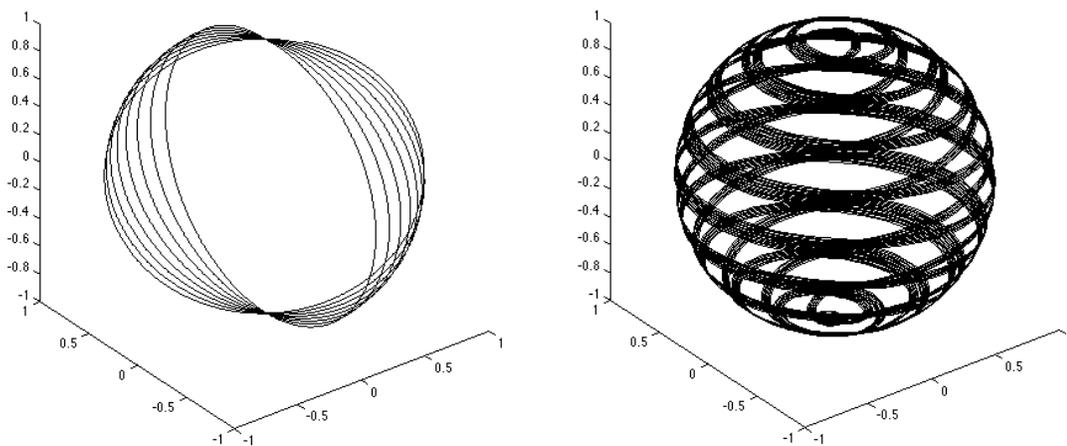


Figure 6.3: Detuned Rabi cycles with  $A = \frac{1}{20}$ ,  $\omega_1 = 0.999$ ,  $\omega_0 = 1$ , and  $t \in [0, 1000]$ .  
 Left: Interaction-picture  $\mathbf{q}(t)$  shown in rotating frame and calculated using RWA.  
 Right: Schrödinger-picture  $\mathbf{r}(t)$  shown in non-rotating frame and calculated without RWA.

## 6.2 Simulated Rabi cycles

Pulse control in a noisy environment requires engineering compromises. Strong pulses make the RWA less accurate. Weak pulses need more time to accomplish the same change in latitude, which can give a qubit enough time to wander badly off course.

A **Rabi cycle experiment** tests how long a pulse can be applied before a qubit spirals out of control. An experimenter chooses  $J$  pulse durations  $\{t_1, \dots, t_J\}$ . For each  $t_j$ , the following steps are repeated many times:

1. Prepare the qubit in its ground state.
2. Apply a resonant pulse.
3. At time  $t_j$ , measure the energy of the qubit.

After many such trials, the number of excited-energy results is used to infer  $P(\text{excited}) = \frac{1}{2}[1 - \bar{z}(t_j)]$ . The inferred  $\bar{z}(t_j)$  values are then plotted as a time series.

For an ideal qubit, Rabi cycles can continue indefinitely. For real qubits, Rabi cycle amplitude typically decays exponentially or approximately so. Attempting to maintain a large Rabi amplitude for as long as possible has become a common test of qubit stability.

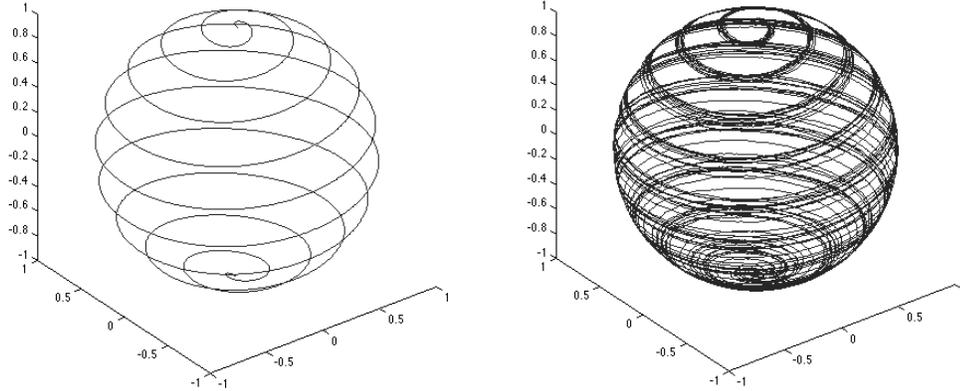


Figure 6.4: Simulated  $\pi$  pulses using `StochasticLinear`.  
Left: Ideal  $\pi$  pulse. Right: 10 trials with  $T_1 = 10000$ ,  $T_2 = 4000$ .

Figure 6.4 shows 10 trials in which a  $\pi$  pulse is applied to a good-but-imperfect qubit in its ground state. For an ideal qubit, the trajectory is a half Rabi cycle. The imperfect trials become visibly non-identical as they spiral toward the South pole. As  $t$  increases, small errors accumulate as each true state wanders away from where it “should” be.

Figure 6.5 shows the mean  $\bar{z}(t)$  coordinate calculated by solving the master equation using the `DrunkenMaster` script. In each case,  $\bar{z}(t)$  follows the damped-oscillator pattern typical of real experiments. For reference,  $e^{-t/T_1}$  and  $e^{-t/T_2}$  are also plotted.

Figure 6.6 shows Monte Carlo simulations of the mean  $\bar{z}(t)$  coordinate using `StochasticLinear`. Each mean state is an average of 250 simulations with 100 timesteps per blink. One full Rabi cycle of one sample state from each simulation is also shown.

All simulations in this section use the linear Bloch model from Chapter 4. The mean field  $\boldsymbol{\mu}$  is given by the pulse control Hamiltonian with  $A = \frac{1}{20}$  and  $\omega_1 = \omega_0 = 1$ :

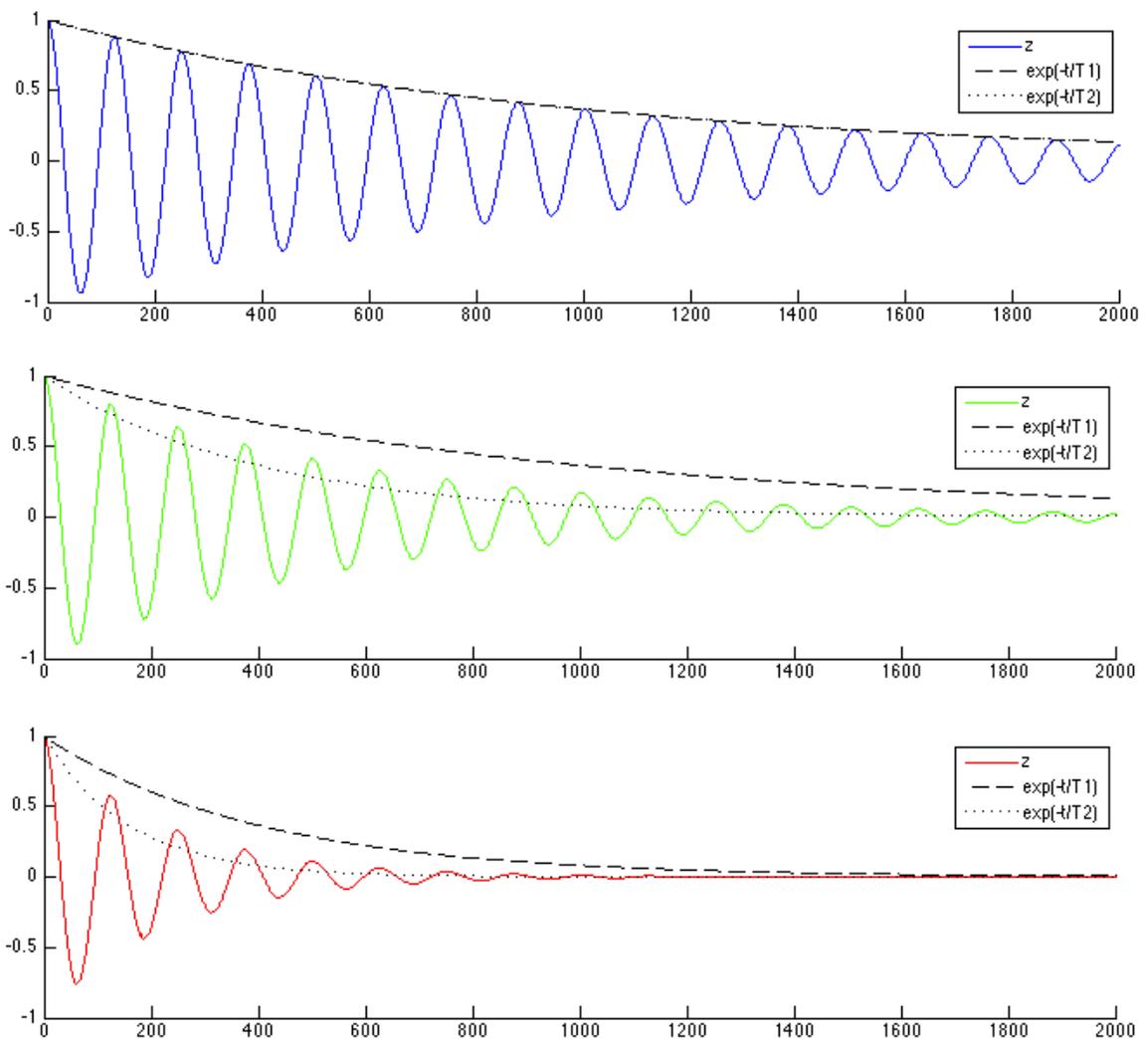
$$\boldsymbol{\mu}(t) = \mathbf{B}(t) = [0.1 \cos(t), 0, 1]^T$$

Each simulation is color-coded by its volatilities  $\nu_1, \nu_z$ . Table 6.1 shows the volatilities and their corresponding  $T_1, T_2$  times calculated from the linear Bloch model.

Table 6.1: Color-coded volatilities for simulations.

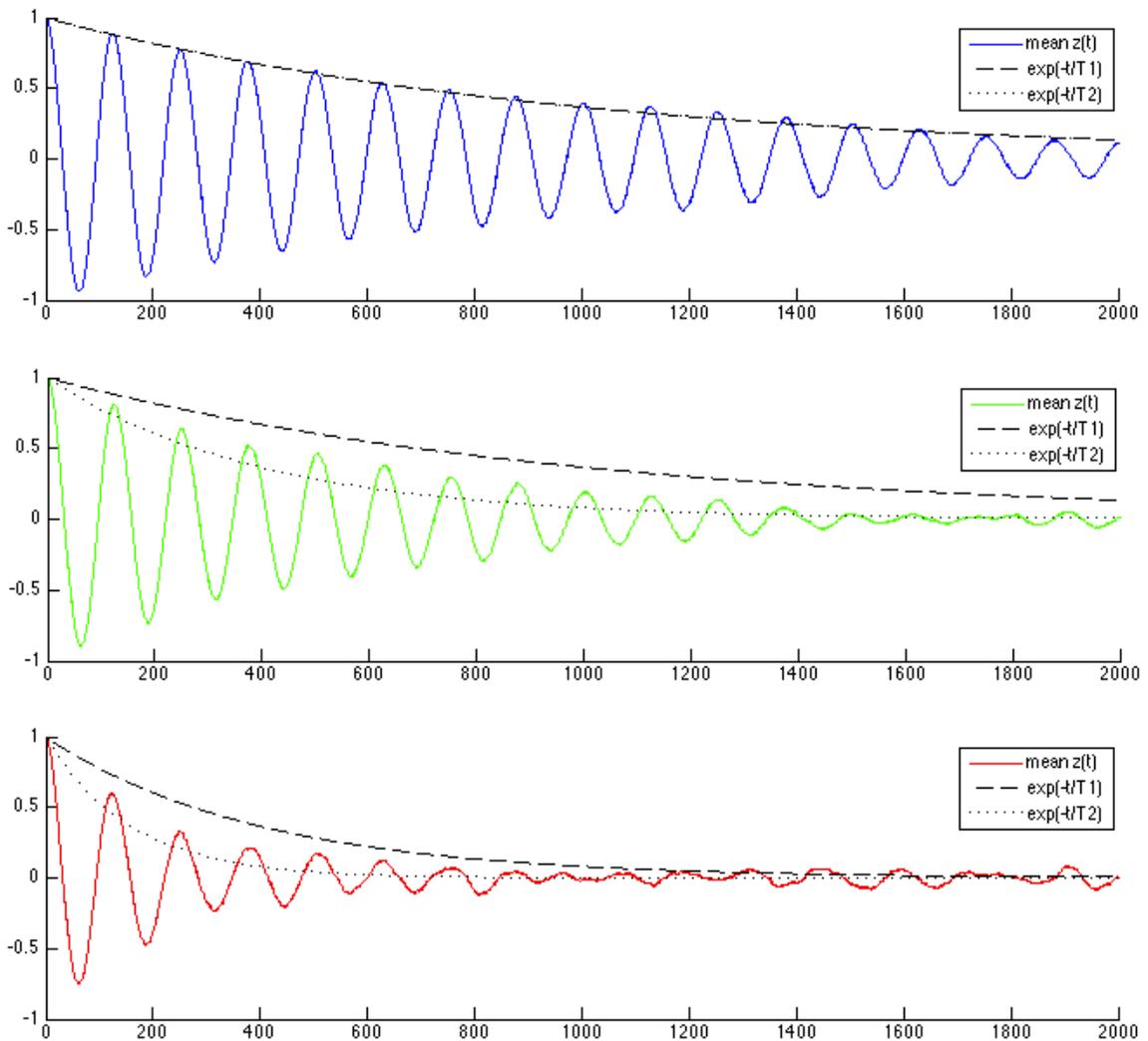
	$\nu_1$	$\nu_z$	$T_1$	$T_2$
blue	.03162	.03162	1000	1000
green	.03162	.06324	1000	400
red	.05	.1	400	160

Figure 6.5: Rabi cycle master equation simulations.



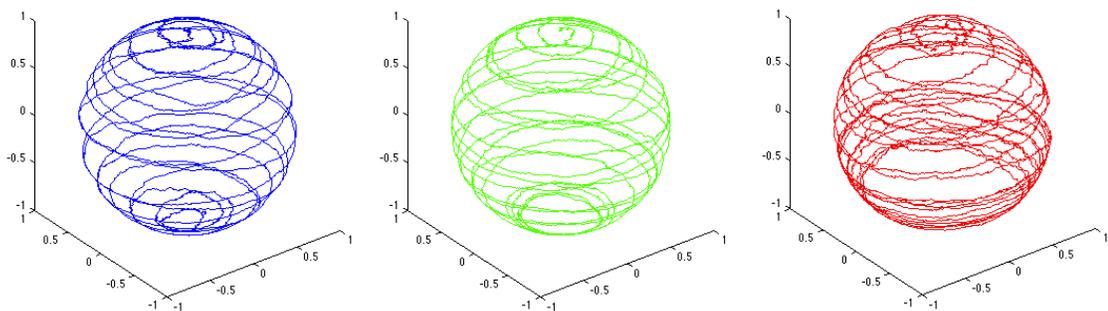
Above: Mean  $\bar{z}(t)$  coordinate found using DrunkenMaster.

Figure 6.6: Rabi cycle Monte Carlo simulations.



Above: Average of 250  $z_t$  coordinates produced by `StochasticLinear`.

Below: Sample state for one full Rabi cycle  $t \in [0, 40\pi]$ .



### 6.3 Simulated Ramsey fringes

**Ramsey fringe** experiments are often used to test how long a qubit can remain in a half-excited state without misbehaving. An experimenter chooses  $J$  waiting times  $\{\tau_1, \dots, \tau_J\}$ . For each  $\tau_j$ , the following steps are repeated many times:<sup>2</sup>

1. Prepare the qubit in its ground state.
2. Apply a  $\pi/2$  pulse to send the qubit's state to the equator.
3. Turn off the pulse and wait  $\tau_j$ .
4. Apply another  $\pi/2$  pulse, then measure the energy of the qubit.

After many such trials, the number of excited-energy results is used to infer  $P(\text{excited}) = \frac{1}{2}[1 - \bar{z}(\tau_j + t_\pi)]$ . The inferred  $\bar{z}(\tau_j + t_\pi)$  values are then plotted as a time series.

The result of the second pulse depends on where the state is at the end of its waiting period. An ideal qubit endlessly circles around the equator with angular frequency  $\omega_0$ . Real qubits tend to wander off course. Averaged over many trials, the mean state  $\bar{\mathbf{r}}(t)$  tends to spiral inward, much like the toy dephasing model in Chapter 4.

Suppose pulses are positive and perfectly resonant:  $A > 0$  and  $\omega_1 = \omega_0 \equiv \omega$ . For an ideal qubit, the first  $\pi/2$  pulse will send the ground state to:

$$\mathbf{r}(t_{\pi/2}) = \hat{S}\mathbf{r}(0) = \begin{bmatrix} \cos(\omega t_{\pi/2}) & 0 & \sin(\omega t_{\pi/2}) \\ -\sin(\omega t_{\pi/2}) & 0 & \cos(\omega t_{\pi/2}) \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \sin(\omega t_{\pi/2}) \\ \cos(\omega t_{\pi/2}) \\ 0 \end{bmatrix}$$

After a waiting time  $\tau$ , the state will be:

$$\mathbf{r}(\tau + t_{\pi/2}) = \hat{T}_0(\tau)\hat{S}\mathbf{r}(0) = \begin{bmatrix} \cos(\omega\tau) & \sin(\omega\tau) & 0 \\ -\sin(\omega\tau) & \cos(\omega\tau) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sin(\omega t_{\pi/2}) \\ \cos(\omega t_{\pi/2}) \\ 0 \end{bmatrix} = \begin{bmatrix} \sin(\omega[\tau + t_{\pi/2}]) \\ \cos(\omega[\tau + t_{\pi/2}]) \\ 0 \end{bmatrix}$$

Applying another  $\pi/2$  pulse sends the state to:

$$\mathbf{r}(\tau + 2t_{\pi/2}) = \hat{S}\hat{T}_0(\tau)\hat{S}\mathbf{r}(0) = \begin{bmatrix} \cos(\omega t_{\pi/2}) \sin(\omega[\tau + t_{\pi/2}]) \\ -\sin(\omega t_{\pi/2}) \sin(\omega[\tau + t_{\pi/2}]) \\ -\cos(\omega[\tau + t_{\pi/2}]) \end{bmatrix}$$

<sup>2</sup>The algorithm shown here is one of several closely-related Ramsey fringe experiments.

The probability of an excited energy result depends sinusoidally on  $\tau$ :

$$P(\text{excited}) = \frac{1}{2} [1 - z(\tau + 2t_{\pi/2})] = \frac{1}{2} + \frac{1}{2} \cos(\omega[\tau + t_{\pi/2}])$$

**Ramsey fringes** are oscillations in the plot of  $P(\text{excited})$  vs.  $\tau$ . As with Rabi cycles, observed Ramsey fringes tend to decay exponentially as  $\tau$  increases. During the waiting period, the mean state spirals exponentially inward with decay rate  $1/T_2$ . For this reason, Ramsey fringe experiments are often used to estimate a qubit's  $T_2$  time.

All simulations in this section use the linear Bloch model from Chapter 4. Positive, perfectly-resonant pulses with  $A = \frac{1}{20}$  and  $\omega_1 = \omega_0 = 1$  are used. The  $\pi/2$  pulse time is then  $t_{\pi/2} = \pi/(2A) = 10\pi$ . The mean field  $\boldsymbol{\mu}(t) = \mathbf{B}(t)$  is a piecewise function:

first pulse	$t \in [0, 10\pi]$	$\mathbf{B}(t) = [0.1 \cos(t), 0, 1]^T$
waiting period	$t \in (10\pi, 10\pi + \tau)$	$\mathbf{B}(t) = [0, 0, 1]^T$
second pulse	$t \in [10\pi + \tau, 20\pi + \tau]$	$\mathbf{B}(t) = [0.1 \cos(t - [\tau + 10\pi]), 0, 1]^T$

Figure 6.7 shows ideal Ramsey trials calculated by `DrunkenMaster` with zero volatility. With these parameters, the final  $z$  coordinate for an ideal trial is  $-\cos(\tau + 10\pi) = -\cos(\tau)$ .

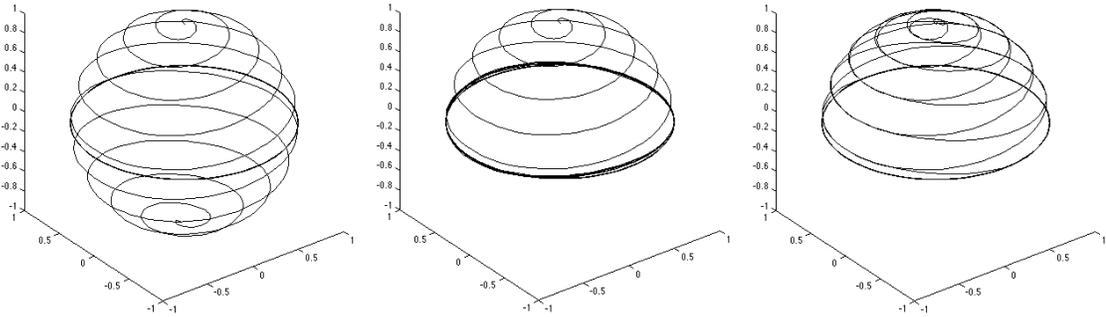
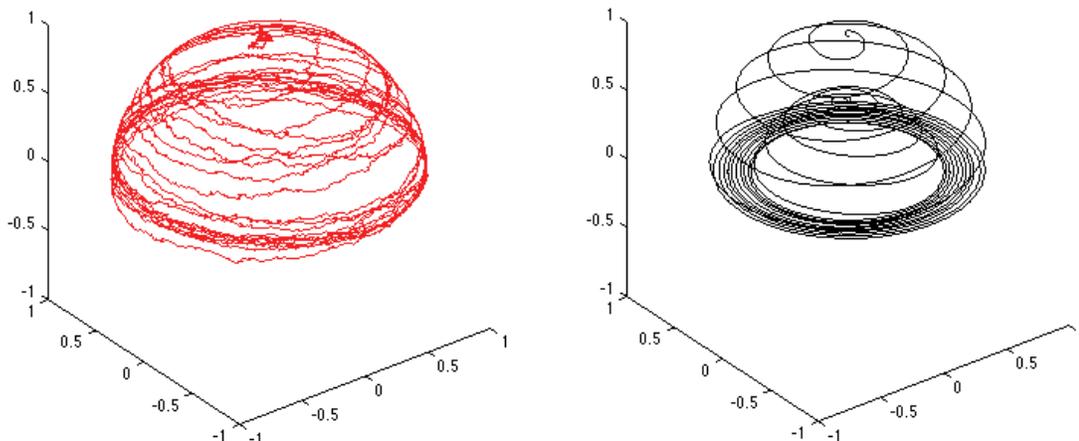


Figure 6.7: Ideal Ramsey trials with  $\tau = 20\pi$ ,  $20.5\pi$ , and  $21\pi$ .

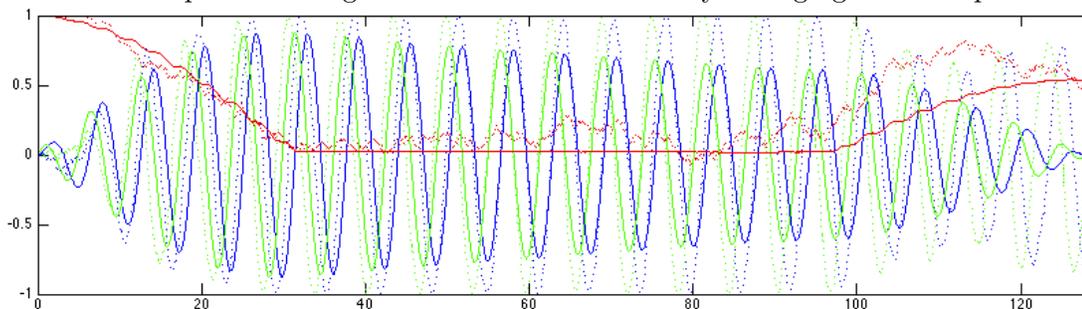
On each trial, a  $\pi/2$  pulse sends the state's Bloch vector from the North pole to the equator. It orbits for a time  $\tau$ , then another  $\pi/2$  pulse is applied.

Figure 6.8 shows the mean state and a sample state from a Monte Carlo simulation of a Ramsey trial with waiting time  $\tau = 21\pi$ . Volatilities were set such that  $T_1 = 400$  and  $T_2 = 160$ . The mean state was estimated by averaging 1000 sample states simulated by `StochasticLinear`. The sample state is typical of simulation results: at first, the qubit approximately obeys its orders and spirals toward the equator. It then orbits unsteadily until the second pulse sends it back North, where it misses the pole by a sizeable margin.

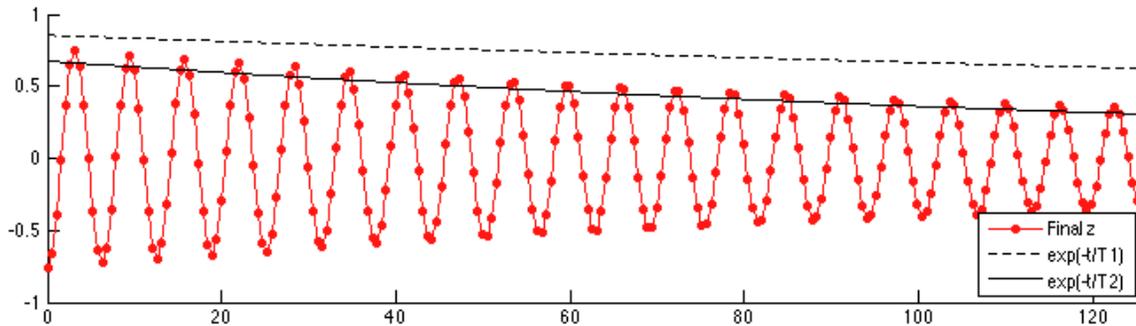
The mean state reaches the equator mostly on course, then steadily dephases inward. The second pulse succeeds in producing an upward spiral, but the mean state's final  $\bar{z}$  coordinate is limited by the many sample states which failed to reach the North pole.

Figure 6.8: Monte Carlo simulation of Ramsey trials with  $\tau = 21\pi$ ,  $T_1 = 400$ ,  $T_2 = 160$ .

Left: One sample state. Right: Mean state estimated by averaging 1000 sample states.

 $x$ ,  $y$ ,  $z$  coordinates of mean state (solid) and one sample state (dotted).

The script `RamseyMaster` simulates a full Ramsey experiment with multiple  $\tau$  values. For each waiting time  $\tau_j$ , it calculates the mean state by numerically integrating the drunken master equation. The final  $\bar{z}$  coordinate is saved, and another simulation is run for the next  $\tau_{j+1}$ . Final  $\bar{z}$  coordinates are plotted versus  $\tau$ . Figure 6.9 shows the results for  $T_1 = 400$ ,  $T_2 = 160$ . For reference,  $\exp(-t_m/T_1)$  and  $\exp(-t_m/T_2)$  are also shown. (Here  $t_m \equiv \tau + t_\pi$  is the total time between the start of a trial and the energy measurement at the end.)

Figure 6.9: `RamseyMaster` simulation with  $T_1 = 400$ ,  $T_2 = 160$ . Vertical axis is final mean  $\bar{z}$  coordinate. Horizontal axis is waiting time  $\tau$ .

## 6.4 Experimental data

Many experiments with various qubit designs have been published in peer-reviewed literature since the early 2000's. As a sample, this section shows excerpts from four papers.

Figure 6.10 is from [88]. The qubit in this experiment is literally a spin- $\frac{1}{2}$  particle:  $|0\rangle$  and  $|1\rangle$  are nuclear spin eigenstates of a phosphorus-31 electron donor in silicon. Radio-frequency pulses are used for qubit control. Readout (i.e. an up-or-down measurement) is performed using electron spin resonance on the donor electron of the  $^{31}\text{P}$  atom. The top two plots in Figure 6.10 are Rabi tests. Each shows spin-flip probability as a function of pulse duration  $t_p$ . The labels  $D^0, D^+$  indicate whether the electron donor was neutral or ionized. The bottom two plots are Ramsey tests for the neutral-donor  $D^0$  and ionized-donor  $D^+$  states. Spin-flip probability is shown as a function of waiting time  $\tau$ .

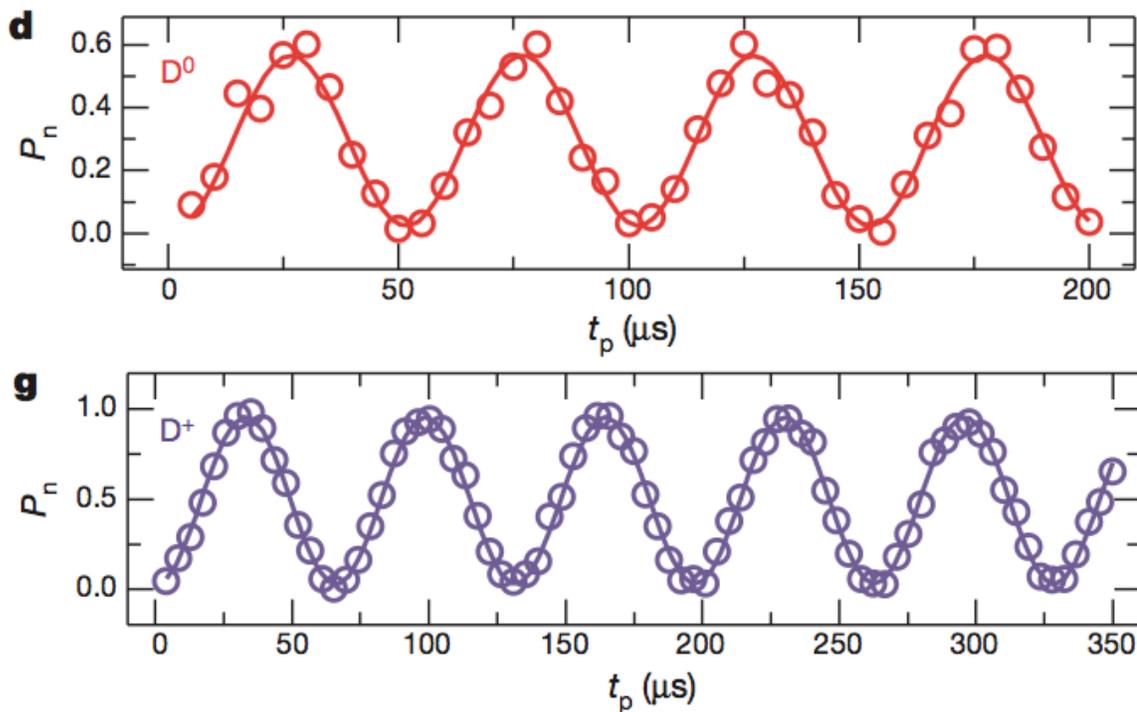
Figure 6.11 is from [89]. Four phase qubits are used, each of which is a SQUID (Superconducting QUantum Interference Device) circuit at mK temperature. Pulse control is accomplished by applying microwave-frequency current to the circuit. The  $|0\rangle$  and  $|1\rangle$  states are *metastable*: they are long-lived scattering resonances, not energy eigenstates. Readout is performed by measuring escape rate, which is much different for  $|0\rangle$  and  $|1\rangle$ . Plots in Figure 6.11 show escape rate  $\Gamma$  as a function of pulse duration  $t$  for a Rabi test.

Figure 6.12 is from [9]. Neutral rubidium atoms are prepared using laser cooling and held in optical dipole traps. Hyperfine atomic states are used as  $|0\rangle$  and  $|1\rangle$ . Pulse control is accomplished using a Raman laser system, and readout is performed by applying radiation pressure with a laser. The readout laser is adjusted such that the  $|1\rangle$  state is forced out of the trap but the  $|0\rangle$  state is not. The top plot in Figure 6.12 shows  $P(\text{ground})$  as a function of pulse duration for a Rabi test. The bottom plot is a Ramsey-like test in which control pulses are detuned so that in a rotating frame, the qubit's Bloch vector drifts out of sync with the pulse.  $P(\text{ground})$  is shown as a function of waiting time  $\tau$ .

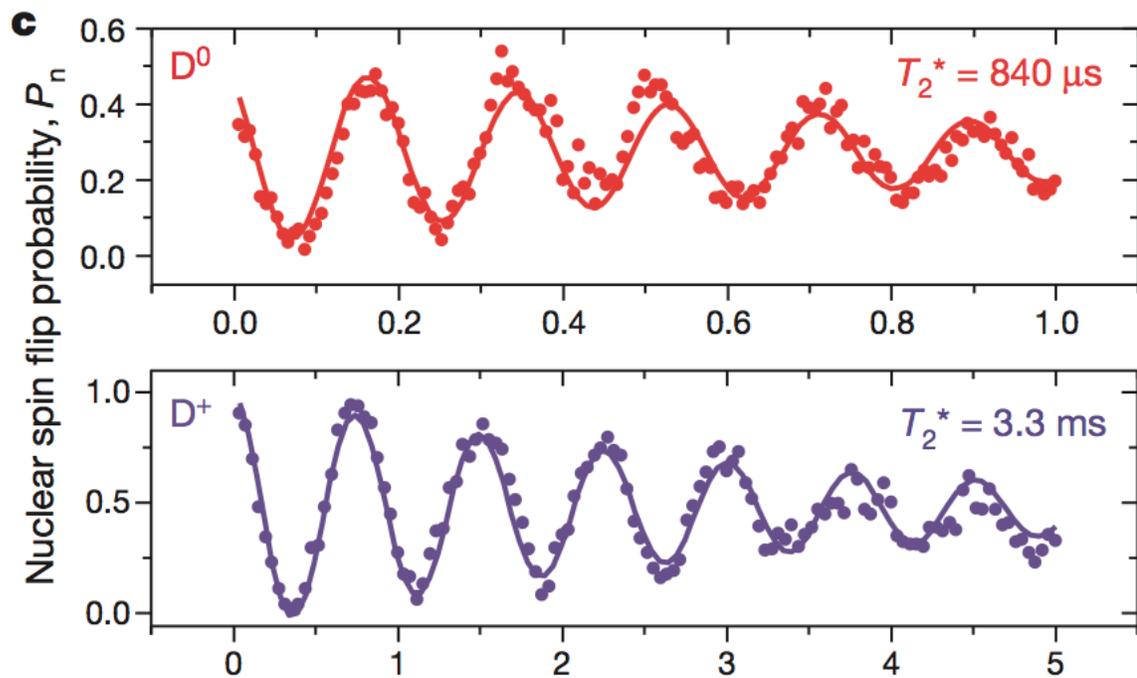
Figure 6.13 is from [90]. Three Josephson junctions in a loop are used as a *flux qubit*. The  $|0\rangle$  state represents a persistent current in one direction, and  $|1\rangle$  represents current in the opposite direction. Pulse control is accomplished by applying microwave-frequency current to the circuit. For readout, a SQUID is attached in such a way that applying a pulse to the SQUID causes it to switch to zero or nonzero voltage depending on the excitation probability of the qubit. The top plots in Figure 6.13 show switching probability as a function of pulse duration for Rabi tests with three different amplitudes. The top-right plot shows observed Rabi frequency as a function of amplitude. The Ramsey interference plot shows switching probability as a function of waiting time  $\tau$ . A *spin-echo* test is also shown.<sup>3</sup>

<sup>3</sup>For brevity, spin-echo tests are omitted from this thesis. See [87] for Hahn's original explanation.

Figure 6.10: Rabi cycles and Ramsey fringes for 2 nuclear spin qubits.



Above: Rabi cycles. Below: Ramsey fringes.



J. J. Pla, et al. "High-fidelity readout and control of a nuclear spin qubit in silicon,"  
*Nature*, vol. 496, Apr 2013.[88]

Figure 6.11: Rabi cycles for 4 phase qubits.

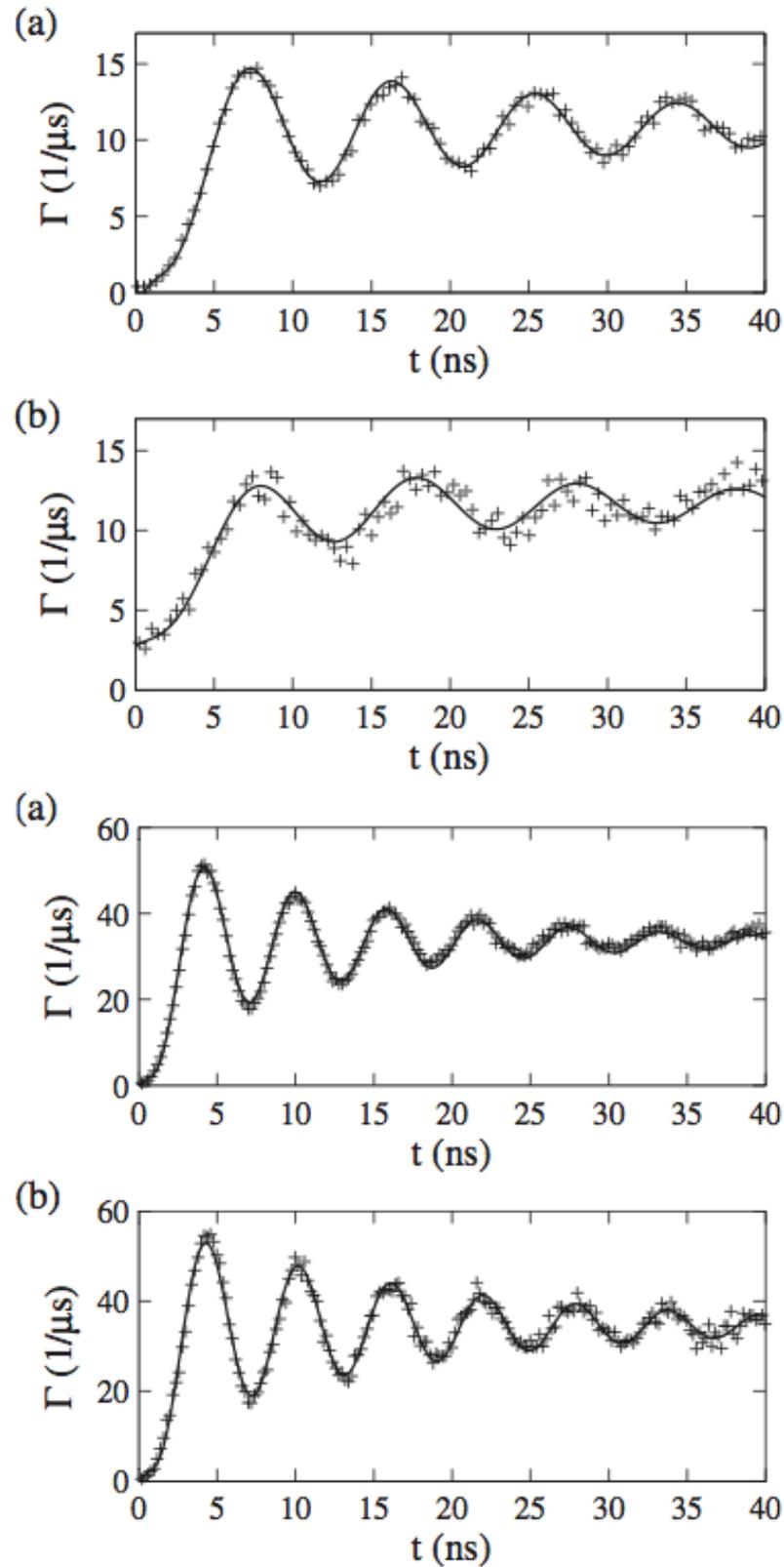
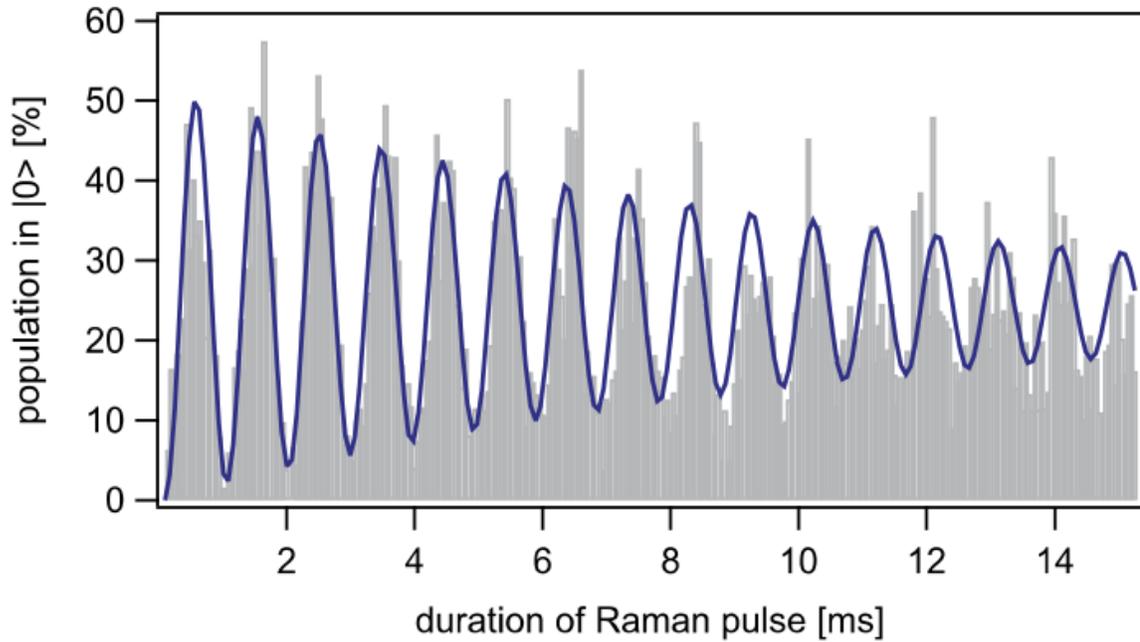
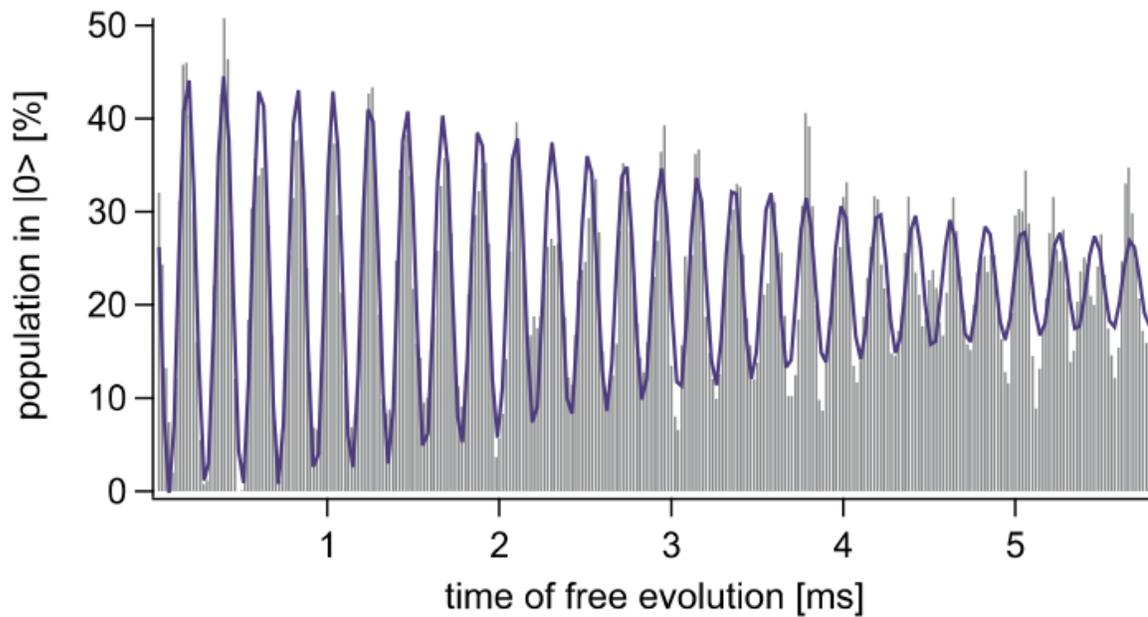


Figure 6.12: Rabi cycles and Ramsey fringes for an atomic qubit.

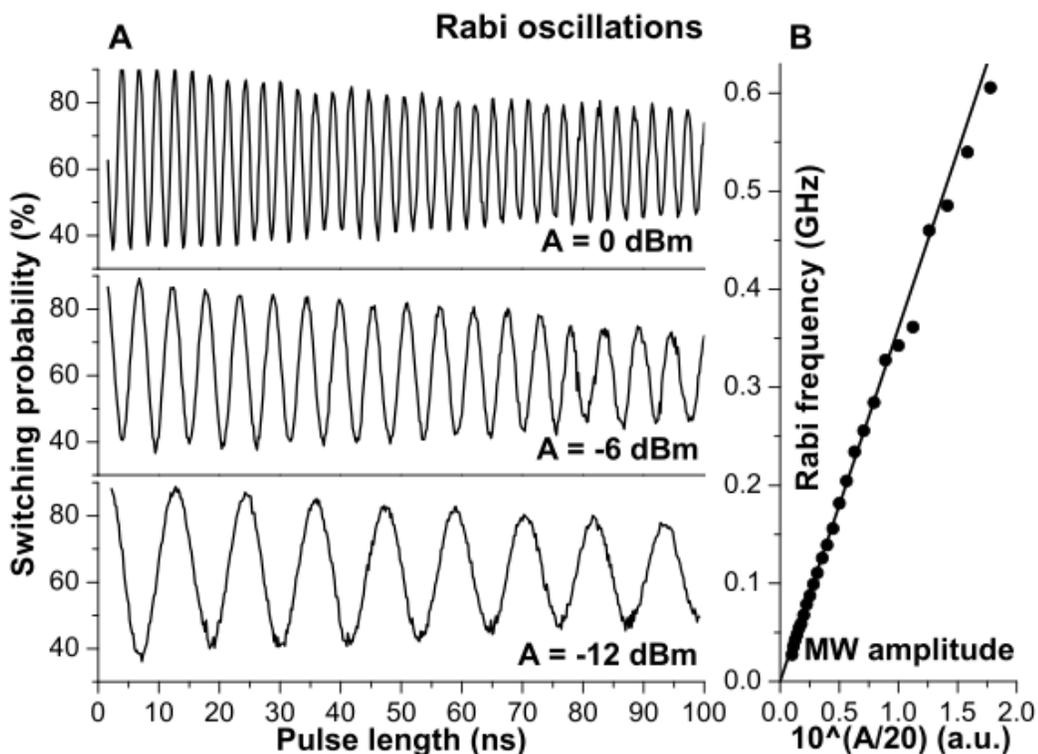


Above: Rabi cycles. Below: Ramsey fringes.

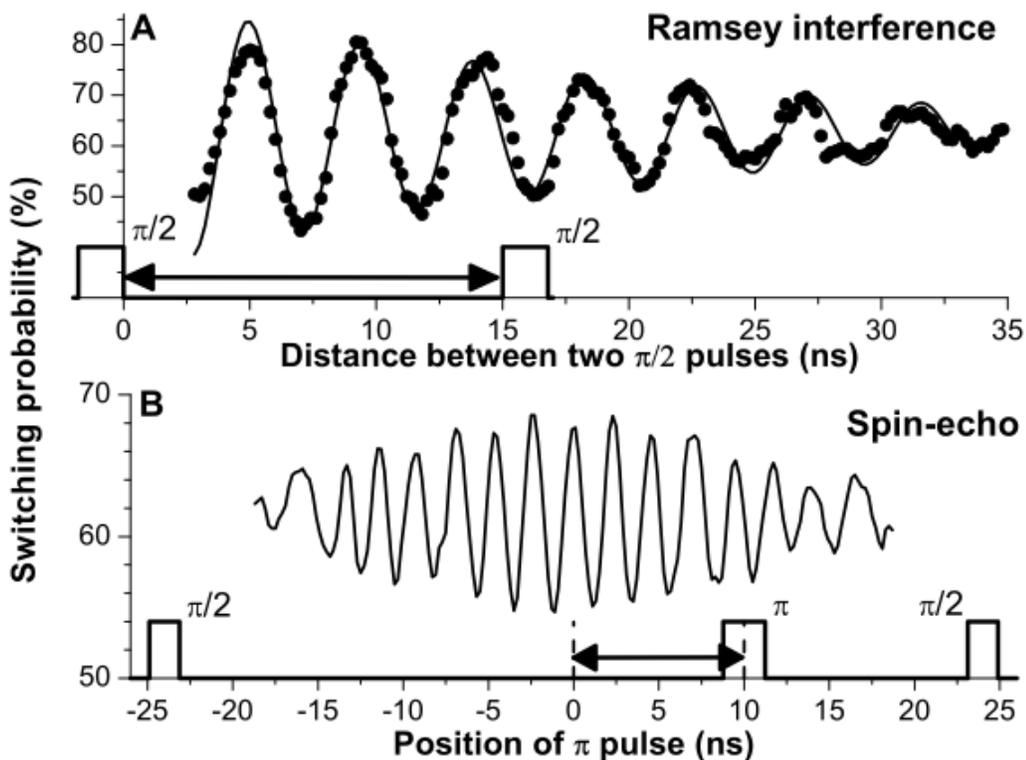


A. Lengwenus, et al. "Coherent manipulation of atomic qubits in optical micropotentials,"  
*Appl. Phys. B*, vol. 86, 2007.[9]

Figure 6.13: Rabi cycles, Ramsey fringes, and spin-echo test for a flux qubit.



Above: Rabi cycles. Below: Ramsey fringes and spin-echo test.



I. Chiorescu, et al. "Coherent quantum dynamics of a superconducting flux qubit," *Science*, vol. 299, Mar 2003.[90]

## 7 CONCLUSIONS

Drunk models provide a quantitative formalism for the effects of non-identical trials during qubit experiments. According to these models, estimated states inferred from experimental data can gain von Neumann entropy even if the true state of the system is always pure.

Many of the practical conclusions of linear drunk models are already common knowledge.<sup>1</sup> The following claims are neither novel nor particularly controversial:

- Any qubit can be represented by a fictional spin- $\frac{1}{2}$  in a magnetic field.
- The Schrödinger equation requires pure states to stay pure.
- Qubit expectation values very often decay exponentially in a way which (at least approximately) resembles the Bloch equation for nuclear magnetic resonance.
- Finding evidence of quantum coherence requires precise experimental controls.
- No experiment can maintain precise control of an unobserved qubit indefinitely.

The main distinctions between drunk models and other descriptions of decoherence are:

1. **No environmental model.** The Hilbert space of possible state vectors is 2-dimensional. Every interaction between a qubit and the rest of the universe is represented as random noise in a  $2 \times 2$  Hamiltonian operator.
2. **Possible nonlinearity.** The mean state of a nonlinear drunk model need not evolve according to a linear master equation.
3. **Pure states only.** Drunk models assume quantum states are always pure. Mixed states are used only as statistical estimators for unknown true states.

The first distinction (no environmental model) may have practical value for the reasons mentioned by Tanimura and Kubo at the beginning of Chapter 4. A generic description of the environment as “random noise” can provide a unified formalism for distinct qubit designs and decoherence processes. Models which do not attempt to describe a qubit’s environment in detail are often mathematically simpler and/or easier to simulate.

---

<sup>1</sup>For a review of what is meant by “common knowledge,” see e.g. [20][47][62][91][5][92] and their cited references. Some of these focus on specific qubit designs, but each provides an overview of recent research.

The second distinction (possible nonlinearity) provides a method for describing experiments in which the estimated state's evolution does not obey the Bloch equation. Solving a nonlinear stochastic differential equation can be extremely difficult, but methods in Chapter 5 and its cited references can be used for Monte Carlo simulations.

The third distinction (pure states only) may be of some theoretical importance. Drunk models predict that pure states *appear* to transform irreversibly into mixed states, but this illusion disappears if one distinguishes between true states, mean states, and estimated states. True states evolve by reversible unitary transformations without exception. Mean and estimated states can gain entropy from accumulation of small errors in control of a system, even if no permanent physical changes occur to the system itself.

## 7.1 A loophole in Loschmidt's paradox

Time evolution of unobserved systems according to the Schrödinger (or Liouville-von Neumann) equation is deterministic, unitary, reversible, and cannot alter von Neumann entropy. This property is not compatible with the maxim that physical systems tend irreversibly toward disorder. One of many versions of the maxim was given by Eddington:

The law that entropy always increases holds, I think, the supreme position among the laws of Nature... If your theory is found to be against the second law of thermodynamics I can give you no hope; there is nothing for it but to collapse in deepest humiliation.[93]

The quantum Loschmidt paradox from Chapter 3 asks: does von Neumann entropy disobey the Second Law of Thermodynamics, or do states evolve in a non-unitary way? According to drunk models, this phrasing of the paradox is a false dichotomy caused by using the word "state" for three distinct mathematical objects:

1. Evolution of the true state of a physical system is deterministic and unitary.
2. Evolution of the mean state of a theoretical model is deterministic and not unitary.
3. The estimated state inferred from experimental data is random and converges to the mean state in the limit of many independent, identically-distributed trials.

Drunk models assume that, at all times, there exists some pure state which is the hypothetical best-possible description of a qubit. Evolution is unitary on each trial, but experimental trials are not identical. A weaker assumption is used: the Hamiltonian is a stochastic process which is *identically distributed* over all trials. Each sample path of represents a Hamiltonian which might occur on an experimental trial, and the mean state is the expectation value of the true state over the set of all possible Hamiltonians and initial conditions.

A consequence of these assumptions is that time evolution is always reversible in principle but often irreversible in practice. If trials of e.g. a Rabi or Ramsey test are very nearly identical, then the inferred  $z(t)$  coordinate will be close to its intended sinusoidal oscillation for sufficiently small values of  $t$ . As  $t$  increases, estimated states tend to decay into mixed states as errors accumulate and each true state wanders off course. Theorists can keep the Schrödinger equation and the Second Law from contradicting each other by assigning the word *entropy* to mean states and estimated states, but never to true states.

### 7.1.1 Anthropomorphic entropy

The assumptions and conclusions of drunk models have much in common with maximum-entropy thermodynamics (MaxEnt), but care should be taken not to conflate the two concepts. Drunk models do not use any thermodynamic postulates and do not purport to explain thermodynamics.<sup>2</sup> The term “macrostate” appears only in Section 2.5, where it is used only for comparison of mean states to thermal-equilibrium density matrices.

The common factor of MaxEnt and drunk models is the use of **anthropomorphic entropy** in a manner similar what Jaynes advocated.[94][95] Drunk models avoid the phrase “entropy of a qubit” because this terminology fails to distinguish between true states of systems, mean states of models, and estimated states of experimental datasets. Mean and estimated states evolve irreversibly even though true states do not. According to Jaynes:

It is possible to maintain the view that the system is at all times in some definite but unknown pure state, which changes because of definite but unknown external forces; the probabilities represent only our ignorance as to the true state. With such an interpretation the expression “irreversible process” represents a semantic confusion; it is not the physical process that is irreversible, but rather our ability to follow it.[95]

Jaynes was describing semiclassical statistical mechanics; drunk models replace “external forces” with “stochastic terms in the Hamiltonian.” In 1965, he added:

From this we see that entropy is an anthropomorphic concept, not only in the well-known statistical sense that it measures the extent of human ignorance as to the microstate. *Even at the purely phenomenological level, entropy is an anthropomorphic concept.* For it is a property, not of the physical system, but of the particular experiments you or I choose to perform on it.[24]

Jaynes attributed the phrase “entropy is an anthropomorphic concept” to Wigner.[24]

---

<sup>2</sup>Nonlinear drunk models *might* be useful for describing the approach of a qubit mean state to thermal equilibrium, but that possibility is not investigated in detail in this thesis.

Drunk models reserve the term *entropy* exclusively for von Neumann entropy of mean states and/or estimated states. This usage is narrower than Jaynes', but similar in spirit. The probabilities assigned to a drunk model's Hamiltonian are not necessarily natural properties of the system being modeled. Rather, they represent that model's inability to precisely describe a system's evolution during an experimental trial. When a drunk-model mean state gains entropy, it is tempting to say "the system is decohering," but it would be more accurate to say "the model is losing its ability to predict the system's true state."

If Jaynes' claim in [95] is taken to its logical extreme, then any physical state must have a "best" representation as a pure state in some (system  $\otimes$  environment) Hilbert space.<sup>3</sup> Drunk models do not contradict this claim – they simply do not attempt to find a system's best representation. Environmental states are ignored, and any system-environment interactions are represented as random noise. In the language of Tanimura and Kubo, a drunk model's stochastic Hamiltonian is merely a model appropriate for the problem rather than an attempt to precisely describe a system and its environment.

### 7.1.2 The shuffle hypothesis

Prior to Jaynes' theories on statistical physics, Eddington considered a somewhat similar approach to resolving time-reversal paradoxes. He hypothesized:

Whenever anything happens which cannot be undone, it is always reducible to the introduction of a random element analogous to that introduced by shuffling.[93]

A literal version of Eddington's shuffle analogy demonstrates why drunk models do not define "the entropy of a system." Suppose a new deck of cards is opened and shown face-up to two players, Allyson and Bob. Both players are asked, "What is the Shannon entropy of the deck?" and both respond "It is zero. The deck is perfectly ordered." Allyson then shuffles the deck many times and places it face-down on the table. During her last shuffle, she accidentally reveals the bottom card to Bob for a moment. For this example, suppose it is  $Q\spadesuit$ . Both players are asked again, "What is the Shannon entropy of the deck?"

Allyson reasons that there are  $52!$  possible states of the deck, and all are approximately equally probable. The entropy of the deck is therefore  $\approx 226$  bits.

$$S \equiv - \sum_{k=1}^{52!} p_k \log_2(p_k) = - \sum_{k=1}^{52!} \frac{1}{52!} \log_2\left(\frac{1}{52!}\right) = \log_2(52!) \approx 226 \text{ bits}$$

Bob says, "Close, but not quite correct. The last card is certainly  $Q\spadesuit$ , so there are  $51!$  equally-probable states. The entropy of the deck is  $\log_2(51!) \approx 220$  bits."

---

<sup>3</sup> $\otimes$  denotes the Hilbert space tensor product, which a colleague facetiously pronounced as *tomato*.

It is intuitively apparent that Bob is cheating in some sense, but that does not falsify his mathematical reasoning. The “entropy of the deck” is less from Bob’s point of view because he has access to more information than Allyson.<sup>4</sup> In an extreme version of the analogy, suppose Allyson’s shuffles were recorded with a high-speed camera. With sufficient patience, anyone with access to the recording could play it back in slow-motion, determine the exact order of the cards, and conclude that the “entropy of the deck” is exactly zero. It appears that the entropy of the deck depends on the mental state of the observer.

This subjectivity can be resolved by using the word “entropy” only to refer to the Shannon entropy of a *random variable* and leaving the phrase “entropy of the deck” undefined. Allyson and Bob each represent the unknown true state of the deck with a random variable. Allyson’s variable  $A$  has a uniform distribution over the set of all  $52!$  possible states. Bob’s  $B$  has a uniform distribution over the set of all  $51!$  possible states for which  $Q\spadesuit$  is at the bottom of the deck. The camera operator’s variable  $C$  assigns probability 1 to the true state and 0 to all other states. The random variables  $A, B, C$  are anthropomorphic in the sense that they are chosen by scientists rather than uniquely determined by Nature. Once these variables are assigned probability distributions, their Shannon entropies are unambiguous: in this case,  $S(A) = \log(52!)$ ,  $S(B) = \log(51!)$ , and  $S(C) = 0$ .

In principle, any shuffle is a permutation, which is a reversible symmetry transformation. In practice, nearly all information about the order of the cards is concealed from all players by a good shuffle. Similarly, drunk models assume states evolve reversibly in ways that are too unpredictable to be known accurately in practice. As with the Allyson’s Choice experiment in Section 3.1, information about the true state of a system is presumed to be *encrypted* by environmental noise, not *destroyed* by it. As with the Zech’s Qubit experiment in Section 3.3 and engineered-decoherence experiments such as [26], observers with a sufficiently accurate recording of a state’s evolution may be able to “rewind” it and conclude its entropy is zero. To avoid subjectivity paradoxes, drunk models define entropy for *density matrices* (or equivalently, points inside the Bloch ball), but never for *physical systems*.

Eddington’s analogy also provides a convenient way to dodge another paradox which sometimes plagues theories of decoherence. If a finite-dimensional system evolves by unitary transformations, then it must eventually return to a state arbitrarily close to its initial state. The formal statement of this property is found in quantum-mechanical versions of the *Poincaré recurrence theorem*. [96][97][98] In terms of Eddington’s shuffle analogy,

There is a ghost of a chance that some day a thoroughly shuffled pack will be found to have come back to the original order. [93]

---

<sup>4</sup>If the deck had been “cut” according to common card-playing practices, then Bob he would only know that  $Q\spadesuit$  is somewhere in the middle of the deck.

Mean states and estimated states are exempt from Poincaré recurrence theorems. (Time evolution of mean and estimated states is not *volume-preserving*, so the theorem does not apply.) The toy dephasing model in Chapter 4 is an example: the trajectory of each true state wanders endlessly around a circle, but the mean state spirals irreversibly inward.

### 7.1.3 Fake decoherence

The term *decoherence* has acquired a variety of different definitions in different contexts. For the models in this thesis, decoherence refers exclusively to von Neumann entropy gain of mean states and estimated states. Researchers often use these terms more broadly. Joos, et al. suggest partitioning the definition of decoherence into three categories called *true*, *false*, and *fake*. Each refers to damping of non-diagonal terms (in some basis) of a density matrix. Abbreviated versions of their definitions from [99] are below:

1. **True decoherence:** The fundamental decoherence mechanism is “pure” entanglement with the environment without any dynamical change of the component states.
2. **False decoherence:** Coherence is trivially lost if one of the required components disappears. An important situation of this kind is represented by relaxation processes.
3. **Fake decoherence:** Decoherence often arises from some averaging process. Two typical situations are noteworthy. The ensemble either consists of members undergoing the same unitary evolution but with different initial states, or an ensemble of identically prepared states subjected to different Hamiltonians is employed. In both cases the fundamental dynamics of a *single* system is unitary, hence there is *no decoherence at all* from a microscopic point of view.

If these definitions are used, then all decoherence predicted by drunk models is fake. Though the word “fake” sounds derogatory, it can be a useful reminder that mean states and estimated states are properties of *models* and *experiments*, not of *physical systems*. Mean states and estimated states are results of averaging processes: mean states are expectation values over an ensemble of initial states and/or different Hamiltonians, and estimated states are sample means of an experimental dataset. Fluctuations are caused by finite-sample-size statistical errors, and dissipation is due entirely to non-identical trials.

As an example of fake decoherence, Joos, et al. describe the experiments [100] and [26] in a way that is similar to the pathwise construction of a drunk qubit in Chapter 2:

It may thus appear that decoherence can also be obtained from “classical perturbations” (kicks) of the quantum system... For “classical noise” the system follows a *unitary* (even though uncontrollable in practice) dynamics; in each individual case it stays in a pure state (that may remain unknown because of an

insufficiently known Hamiltonian)... In contrast, decoherence leads deterministically to an entangled state that has quite different properties. “Noise” models are in fact only used in situations where this difference cannot be observed.[99]

Zurek also emphasizes that fake decoherence is reversible in principle, even if it is irreversible in practice. His description of the engineered-decoherence experiments [26][101][102] is conceptually similar to drunk models and Jaynes’ view of entropy:

Following a particular realization of the noise, the state of the system is still pure. Nevertheless, an ensemble average over many noise realizations is represented by the density matrix that follows an appropriate master equation. Thus, as Wineland, Monroe, and their colleagues note, decoherence simulated by classical noise could be in each individual case – for each realization – reversed by simply measuring the corresponding time-dependent noise either beforehand or afterwards, and then applying the appropriate unitary transformation to the state of the system. By contrast, in the case of entangling interactions, two measurements, one preparing the environment before the interaction with the environment, the other following it, would be the least required for a chance of undoing the effect of decoherence.[103]

Mean states are ensemble averages over all possible noise realizations, which is similar to Zurek’s description. But drunk models do not necessarily assume that all noise is caused by random classical fields. The Zech’s Qubit thought experiment in Section 3.3 uses a random classical field as a noise source, but this is only an example chosen for the sake of simplicity. The stochastic process  $\mathbf{B}_t$  is merely a convenient choice of parameters for an erratic Hamiltonian and need not be an actual classical field.

## 7.2 Comparison with other theories

A complete list of every theory of decoherence – or even those directly relevant to drunk models – would be far too large to fit in these pages. This section briefly describes the relation of the models in this thesis to two broad classes of decoherence theories: those which attempt to accurately represent a system’s environment, and those which replace the Schrödinger equation with a master equation permitting non-unitary evolution. The practical motivation for all of these theories is, in Zurek’s words,

In the absence of the ideal – a completely isolated absolutely perfect quantum computer, something easy for a theorist to imagine but impossible to attain in the laboratory – one must deal with imperfect hardware leaking some of its information to the environment.[103]

---

Drunk models fall into the non-unitary category. Environmental details are deliberately ignored for the sake of calculational simplicity. A system’s true state and true Hamiltonian are presumed to be unattainable mathematical idealizations. Their purpose is to provide deliberately-imperfect descriptions of experiments for which accurate representations of a qubit’s environment are impossible, impractical, or simply too difficult to solve. In these situations, drunk models are a consolation prize; a random model is much better than no model. In terms of Eddington’s shuffle analogy: if we can’t predict the order of a deck of cards, we can still represent it as a random variable and calculate odds.

For some experiments, a gambler’s attitude may be too pessimistic. One well-known alternative is to represent a system’s environment as a “bath” of infinitely-many harmonic oscillators. Feynman and Vernon proposed a harmonic-bath model in a 1963 paper based on part of Vernon’s PhD thesis.[104] In the 1980’s, Caldeira and Leggett used Feynman-Vernon *influence functionals* to find an exactly soluble model of environmental damping of two Gaussian wave packets in a harmonic potential.[105] Caldeira’s PhD thesis also used Feynman-Vernon methods to derive a density matrix for Brownian motion.[106] Joos, et al. mention other examples: [107] uses a harmonic bath to model dissipative quantum transport, and [108][109] describe Stern-Gerlach and EPR-type experiments using Caldeira-Leggett methods. A cite-search of [104] and [106] reveals hundreds of other examples in which specific environmental models are used to describe dissipative quantum systems.

If it is possible to explicitly describe the details of a system’s environment, and it is practical to solve or approximate the resulting model to sufficient precision, then drunk models may appear unnecessary. In the author’s opinion, this is a perfectly reasonable conclusion. Why pretend a shuffle is random when one has a detailed model of the shuffling mechanism?

Explicit environmental models necessarily require extra degrees of freedom – in some cases, infinitely many. For many experiments, an explicit model may be impractical to derive and/or too difficult to solve accurately. An alternative is provided by master-equation methods. The most well-known of these are the *Kossakowski-Lindblad master equations* (KLM) developed in the 1970’s. Like drunk models, KLM equations model the influence of a system’s environment without attempting to model the environment itself in detail. The best description of a system and its environment is presumed to be a pure state in some larger (system  $\otimes$  environment) Hilbert space, and the environment is “averaged over” by performing a *partial trace*. The next subsection briefly reviews KLM equations and their relationship to the drunken master equation from Section 4.1.

### 7.2.1 Kossakowski-Lindblad master equations

To define a **Kossakowski-Lindblad master equation** for an  $N$ -level system, choose  $N^2 - 1$  linear operators  $\{\hat{L}_m\}$  which are traceless and Hilbert-Schmidt orthogonal:

$$\text{Tr}[\hat{L}_m] = 0 \qquad \text{Tr}[\hat{L}_n \hat{L}_m] = \delta_{nm}$$

Choose an  $(N^2 - 1) \times (N^2 - 1)$  complex matrix  $\hat{C}$  whose eigenvalues are non-negative real, and let  $\{c_{nm}\}$  denote its elements. Time evolution of a density matrix  $\bar{\rho}$  is given by:[52]

$$\hbar \frac{d}{dt} \bar{\rho} = -i[\hat{H}, \bar{\rho}] + \frac{1}{2} \sum_{n=1}^{N^2-1} \sum_{m=1}^{N^2-1} c_{nm} \left( [\hat{L}_n, \bar{\rho} \hat{L}_m^\dagger] + [\hat{L}_n \bar{\rho}, \hat{L}_m^\dagger] \right) \quad (\text{KLM})$$

which is often written in an alternate form:

$$\hbar \frac{d}{dt} \bar{\rho} = -i[\hat{H}, \bar{\rho}] + \sum_{n=1}^{N^2-1} \sum_{m=1}^{N^2-1} c_{nm} \left( \hat{L}_n \bar{\rho} \hat{L}_m^\dagger - \frac{1}{2} \{ \hat{L}_m^\dagger \hat{L}_n, \bar{\rho} \} \right)$$

Here  $\{, \}$  is the matrix anticommutator  $\{\hat{A}, \hat{B}\} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}$ . Equations of the form (KLM) can predict non-unitary evolution of  $\bar{\rho}$ . According to Lindblad, their purpose is:

The dynamics of a finite closed quantum system is conventionally represented by a one-parameter group of unitary transformations in Hilbert space. This formalism makes it difficult to describe irreversible processes like the decay of unstable particles, approach to thermodynamic equilibrium and measurement processes.

It seems that the only possibility of introducing an irreversible behavior in a finite system is to avoid the unitary time development altogether by considering non-Hamiltonian systems. One way of doing this is by postulating an interaction of the considered system  $S$  with an external system  $R$  like a heat bath or a measuring instrument... A different physical interpretation with the same mathematical structure is to consider  $S$  as a limited set of (macroscopic) degrees of freedom of a large system  $S + R$  and  $R$  as the uncontrolled (microscopic) degrees of freedom. If the reservoir  $R$  is supposed to be finite (but large) then the development of the system  $S + R$  may be given by a unitary group of transformations. The partial state of  $S$  then suffers a time development which is not given by a unitary transformation in general.[110]

The drunken master equation derived in Section 4.1 is a KLM equation. To see the relationship, choose rescaled Pauli matrices as the operators  $\{\hat{L}_m\}$ :<sup>5</sup>

$$\hat{L}_1 = \frac{1}{\sqrt{2}}\hat{\sigma}_x \quad \hat{L}_2 = \frac{1}{\sqrt{2}}\hat{\sigma}_y \quad \hat{L}_3 = \frac{1}{\sqrt{2}}\hat{\sigma}_z$$

After some straightforward but tedious algebra (or symbolic *Mathematica* computation), the resulting KLM equation can be written in terms of Pauli coordinates:

$$\frac{d}{dt}\bar{\mathbf{r}} = \mathbf{H} \times \bar{\mathbf{r}} + \hat{M}\bar{\mathbf{r}} + \mathbf{b}$$

where the matrix  $\hat{M}$  and column vector  $\mathbf{b}$  are:

$$\hat{M} \equiv \begin{bmatrix} -c_{22} - c_{33} & \frac{1}{2}(c_{12} + c_{21}) & \frac{1}{2}(c_{13} + c_{31}) \\ \frac{1}{2}(c_{12} + c_{21}) & -c_{11} - c_{33} & \frac{1}{2}(c_{23} + c_{32}) \\ \frac{1}{2}(c_{13} + c_{31}) & \frac{1}{2}(c_{23} + c_{32}) & -c_{11} - c_{22} \end{bmatrix} \quad \mathbf{b} \equiv \begin{bmatrix} i(c_{23} - c_{32}) \\ i(c_{31} - c_{13}) \\ i(c_{12} - c_{21}) \end{bmatrix}$$

Suppose the coefficient matrix  $\hat{C}$  is real symmetric. Then  $\mathbf{b} = \mathbf{0}$ , and:

$$\hat{M} = \begin{bmatrix} -c_{22} - c_{33} & c_{12} & c_{13} \\ c_{21} & -c_{11} - c_{33} & c_{23} \\ c_{31} & c_{32} & -c_{11} - c_{22} \end{bmatrix} = \hat{C} - \text{Tr}[\hat{C}]\hat{1}$$

Recall from Section 4.1 that any qubit drunken master equation can be written

$$\frac{d}{dt}\bar{\mathbf{r}} = -\boldsymbol{\mu} \times \bar{\mathbf{r}} + \frac{1}{2}(\hat{\Sigma}^T \hat{\Sigma} - \text{Tr}[\hat{\Sigma}^T \hat{\Sigma}]\hat{1})\bar{\mathbf{r}}$$

Any qubit drunken master equation is a KLM equation whose  $\{\hat{L}_m\}$  operators are rescaled Pauli matrices and whose effective  $\mathbf{H}$  and coefficient matrix  $\hat{C}$  are:

$$\mathbf{H} = -\boldsymbol{\mu} \quad \hat{C} = \frac{1}{2}\hat{\Sigma}^T \hat{\Sigma}$$

The derivation of the drunken master equation is very different from the methods used by Kossakowski and Lindblad, but the strategies are not entirely antithetical. From a physical perspective, each assumes that a qubit is described by averaging over some larger system which is a nuisance to model accurately. But while all qubit drunken master equations are KLM equations, the converse is false. In particular, any Bloch equation with a nonzero steady-state solution is a KLM equation but not a drunken master equation.

---

<sup>5</sup>Rescaling is necessary to comply with the convention  $\text{Tr}[\hat{L}_n \hat{L}_m] = \delta_{nm}$ . The annoying factor of 2 has returned from Chapter 1 to annoy us once again!

The Bloch equation for a mean state  $\bar{\mathbf{r}} = [\bar{x}, \bar{y}, \bar{z}]^T$  is:

$$\begin{aligned}\frac{d}{dt}\bar{x} &= -(\boldsymbol{\mu} \times \bar{\mathbf{r}})_x - \frac{\bar{x}}{T_2} \\ \frac{d}{dt}\bar{y} &= -(\boldsymbol{\mu} \times \bar{\mathbf{r}})_y - \frac{\bar{y}}{T_2} \\ \frac{d}{dt}\bar{z} &= -(\boldsymbol{\mu} \times \bar{\mathbf{r}})_z - \frac{\bar{z} - z_\infty}{T_1}\end{aligned}$$

where  $z_\infty$  denotes the steady-state  $\bar{z}$  coordinate. If the steady-state value of  $\bar{\mathbf{r}}$  is the infinite-temperature canonical mixed state, then  $z_\infty = 0$ . In this case, the Bloch equation is the drunken master equation derived from the linear Bloch model in Section 4.5. But if  $z_\infty \neq 0$ , then the Bloch equation is not a drunken master equation.

### 7.3 Future research

This section briefly suggests possible generalizations and improvements.

#### Nonlinear drunk models

A nonlinear drunk model was introduced in Section 4.6, but it was not developed in detail. The reason is simple: solving nonlinear stochastic differential equations is difficult, and the master equation derivation in Section 4.1 does not work for nonlinear models. Readers who are interested in nonlinear models are encouraged to use and/or modify the MATLAB script `StochasticNonlinear` in Appendix C for Monte Carlo simulations.

#### $N$ -level systems

One obvious limitation of the models in this thesis is that each describes a single qubit. Wherever possible, definitions and assumptions have been designed to assist any future attempts at generalizing to arbitrary finite-dimensional systems. Appendix B outlines methods for generalizing Pauli coordinates to higher dimensions. Fano's 1957 review and a related paper by Weigert from 1999 may be especially helpful for this task.[2][12] Fox's prior work on stochastic Hamiltonians is also very relevant.[50][51][49]

The distinctions between drunken master equations and KLM equations may be more prominent for  $N$ -level systems with  $N > 2$ . Gorini and Kossakowski noted that:

...unless  $N = 2$ , not all center-preserving completely positive dynamical maps of an  $N$ -level system can be obtained as convex combinations of unitary transformations.[53]

The mean state of any drunk model is, by definition, a convex combination of pure states which are each unitary transformations of an initial state. This difference may be a major incompatibility between drunk models and KLM equations in higher dimensions.

### Systems with memory

As was briefly mentioned in Section 4.6, all the models in this thesis are Markovian. From a physical point of view, the Markov property can be heuristically stated as “a system might remember where it is, but not how it got there.” Markovian processes may be inadequate for modeling systems which “remember” prior states. As early as 1957, Jaynes had warned about the difficulty of non-Markovian stochastic models:

In the case of a system perturbed by random fluctuation fields, the density matrix cannot satisfy any differential equation because  $\dot{\rho}(t)$  does not depend only on  $\rho(t)$ , but also on past conditions. The rigorous theory involves stochastic equations of the type  $\rho(t) = G(t, 0)\hat{\rho}(0)$ , where the operator  $G$  is a functional of conditions during the entire interval ( $0 \rightarrow t$ ). Therefore a general theory of irreversible processes cannot be based on differential rate equations corresponding to time-proportional transition probabilities. However, such equations often represent useful approximations.[95]

The mean state of a linear drunk model does not depend on past conditions, and it does indeed satisfy a differential equation (the drunken master equation). This luxury is available because  $\dot{\rho}(t)$  does not depend on  $\rho(s)$  unless  $s = t$ .

### Colored and/or non-Gaussian noise

As mentioned in the introduction to Chapter 4, all of the stochastic differential equations in this thesis are either Itô or Stratonovich-Fisk equations involving Gaussian white noise (GWN). Here *white noise* means a stochastic process whose power spectral density is constant over all frequencies, and *Gaussian noise* means a stochastic process whose increments are normally distributed. The exclusive use of GWN in this thesis commits two common statistical misdeeds: assuming random variables are Gaussian, and assuming increments of a noise process are independent. This decision was made primarily for convenience, as most of the relevant mathematics can be found in textbooks such as [46][57][58] and [59].

Two alternatives which may be of interest are *fractional Brownian motion* and *functional Itô calculus*. Fractional Brownian motion was developed by Mandelbrot and van Ness to describe colored Gaussian noise.[111] Functional Itô calculus was developed by Dupire to generalize Itô calculus for path-dependent functions.[112]

## A STOCHASTIC CALCULUS

This appendix is intended as a brief overview of the methods and terminology used in Chapter 4. For thorough definitions and explanations, see textbooks such as [46][57][58][59].

For many practical purposes, a **stochastic process** can be thought of as a random variable which changes in time. A common example is the **simple random walk**. Suppose a gambler bets \$1, with even-money odds, that a fair coin flip will land heads. Represent the result of each toss with a random variable  $X_t$  which has two outcomes 1 and -1, each of which occurs with probability  $\frac{1}{2}$ . Let  $S_t$  denote the gambler's profit after  $t$  tosses:

$$S_t \equiv \sum_{j=1}^t X_j$$

It is a common abuse of notation to say “the stochastic process  $S_t$ ,” though it would be more accurate to call the sequence of random variables  $\{S_0, S_1, S_2, \dots\}$  a stochastic process. The *Wiener process*  $W_t$  is, roughly speaking,  $S_t$  in the limit that the coins are flipped very rapidly with very small bet sizes designed to keep the variance per hour unchanged.

As a heuristic for stochastic differential equations, consider a difference equation:

$$Y_{t+1} = Y_t + 3 + X_t \quad Y_0 = 0$$

where  $X_t$  is defined as before. If there were no noise term  $X_t$ , then  $Y_t$  would evolve deterministically and the solution would be  $Y_t = 3t$ . Intuitively, one might correctly guess that  $3t$  is the expected value of  $Y_t$  because each toss has expected value zero. Now suppose the amplitude of the noise is increased so that the difference equation becomes

$$Y_{t+1} = Y_t + 3 + 10X_t \quad Y_0 = 0$$

Figure A.1 shows two Monte Carlo simulations consisting of 42 trials each of the “quiet” and “loud” difference equations. The “loud” version is far more volatile, but its expected value still drifts upward at the same rate of \$3 per toss.

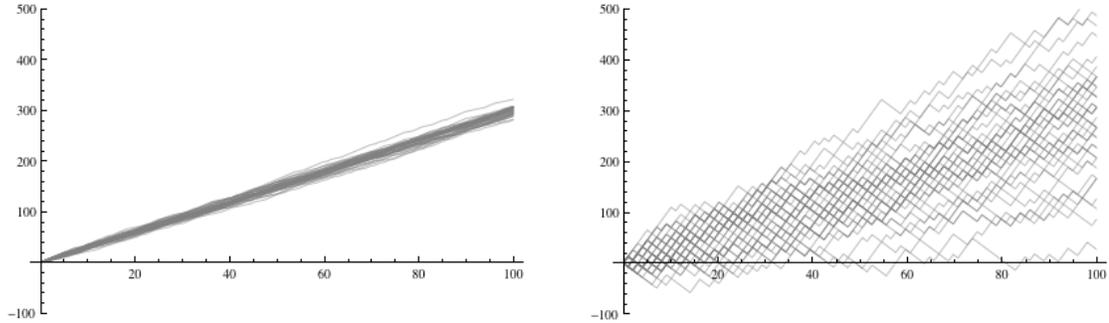


Figure A.1: 42 simulated sample paths for  $Y_t$  as  $t$  advances from 0 to 100.  
 Left: “quiet”  $Y_{t+1} = Y_t + 3 + X_t$ . Right: “loud”  $Y_{t+1} = Y_t + 3 + 10X_t$ .

## A.1 Itô and Stratonovich-Fisk calculus

An **increment** of a stochastic process  $X$  over a time interval  $[s, t]$  is a random variable  $X_t - X_s$ . The following properties define **the Wiener process**  $W$ :

0.  $W_0 = 0$
1. Each increment  $W_t - W_s$  is normally-distributed with mean 0 and variance  $|t - s|$ .
2. Increments over non-overlapping intervals are independent, e.g.  $W_2 - W_1$  is independent of  $W_1 - W_0$ .
3.  $W$  is almost surely continuous at any  $t$ .

In general, one must clearly specify whether a stochastic differential equation is meant to be solved by Itô or Stratonovich-Fisk rules. For drunk models, the ability to switch between these two formalisms is very useful. The essential ideas are:

- $dW$  is “an infinitesimal increment of  $W$ ,” but this is ill-defined by ordinary calculus.
- Itô and Stratonovich-Fisk (SF) stochastic calculi are two distinct formalisms for sensibly writing  $W_t = \int_0^t dW$ .
- The *Wong-Zakai correction* converts Itô SDEs to SF SDEs and vice versa.

Special rules of calculus are needed because  $W_t$  is almost surely *not* differentiable at any  $t$ . Informally,  $W_t$  is “infinitely squiggly” in the sense that one can zoom in very close on a plot of  $W_t$  without finding a good linear approximation.<sup>1</sup> However, it is still possible to define  $\int dW$  in a way similar to Riemann-Stieltjes (RS) integration. Suppose  $f, g$  are two

<sup>1</sup>More precisely,  $W_t$  has nonzero *quadratic variation*.

functions of time. The RS integral  $\int f(t)dg(t)$  is:

$$\int_{t_0}^{t_J} f(t) dg(t) \equiv \lim_{J \rightarrow \infty} \sum_{j=0}^{J-1} f(t_j) [g(t_{j+1}) - g(t_j)]$$

From a numerical point of view, the time interval  $[t_0, t_J]$  is broken into  $J$  timesteps with  $J+1$  endpoints  $\{t_0, \dots, t_J\}$ . On each timestep, the computer multiplies  $f(t_j)$  by the change in  $g$  during the step. If timesteps are sufficiently short, then the choice of when to evaluate  $f$  during each step becomes arbitrary. For It $\bar{o}$  and Stratonovich-Fisk integrals, the choice of when to evaluate  $f$  is *not* arbitrary. It $\bar{o}$  integrals evaluate  $f(t)$  at the beginning of each timestep, and SF integrals use the midpoint.

$$\int_{t_0}^{t_J} f(t) dW \equiv \lim_{J \rightarrow \infty} \sum_{j=0}^{J-1} f(t_?) [W_{t_{j+1}} - W_{t_j}] \quad \begin{array}{l} \text{It}\bar{o}: t_? = t_j \\ \text{SF}: t_? = \frac{1}{2}(t_j + t_{j+1}) \end{array}$$

It $\bar{o}$  and SF integrals are the limits (in probability) of these sums for small timesteps. The definition can also be extended for cases in which the integrand is a stochastic process.<sup>2</sup>

### Stochastic differential equations

Consider a one-dimensional SDE with one noise source:

$$dX_t = g(t, X_t)dt + h(t, X_t)dW \quad (\text{A.1})$$

The functions  $g$  and  $h$  are **drift** and **diffusion** terms. (A.1) is simplified notation for

$$X_T = X_0 + \int_0^T g(t, X_t)dt + \int_0^T h(t, X_t)dW$$

where the second integral is a stochastic integral. Equation (A.1) is ambiguous until one specifies *which stochastic calculus to use*. A common convention is to use  $dW$  to denote It $\bar{o}$  calculus and  $\circ dW$  for SF calculus. Two important distinctions are:

- The expectation value  $E[\int_0^T f(t, X_t) dW]$  is 0, but  $E[\int_0^T f(t, X_t) \circ dW]$  need not be.
- The Chain Rule from ordinary calculus works for SF calculus, but it fails for It $\bar{o}$  calculus and must be replaced by *It $\bar{o}$ 's Lemma*.

<sup>2</sup>It $\bar{o}$  integrals  $\int X_t dW$  are well-defined only if  $X_t$  is *adapted to the natural filtration of  $W_t$* . Informally, this means “the integrand cannot see into the future,” i.e.  $X_t$  probabilities can depend on events that occurred at or before time  $t$ , but not after  $t$ . For measure-theoretic definitions, see e.g. [57] or [58].

### Itô's Lemma

Suppose  $f$  is some function of  $t$  and  $x$  with well-behaved derivatives  $\dot{f} \equiv \partial_t f$ ,  $f' \equiv \partial_x f$ , and  $f'' \equiv \partial_x^2 f$ . Then  $f(t, W_t)$  solves the following SDEs:

$$\begin{aligned} \text{Itô's Lemma:} \quad & df(t, W_t) = \dot{f}(t, W_t)dt + f'(t, W_t)dW + \frac{1}{2}f''(t, W_t)dt \\ \text{SF Chain Rule:} \quad & df(t, W_t) = \dot{f}(t, W_t)dt + f'(t, W_t) \circ dW \end{aligned}$$

The SF Chain Rule conveniently resembles the ordinary Chain Rule if  $W_t$  were a differentiable function of time. For Itô SDEs, change-of-variables formulas must be calculated using Itô's Lemma instead. This curiosity can be a dangerous nuisance when solving Itô SDEs on curved manifolds. However, expectation values of Itô SDEs are much easier to calculate. One can simply erase the  $dW$  terms:

$$E[f(t_J, W_{t_J})] = f(t_0, W_{t_0}) + \int_{t_0}^{t_J} \left( \dot{f}(t, W_t) + \frac{1}{2}f''(t, W_t) \right) dt + \int_{t_0}^{t_J} \cancel{f'(t, W_t) dW}$$

This magic trick follows from the *martingale* property of Itô integrals. Roughly speaking, martingales are mathematical representations of “fair games” whose expectation values are zero no matter what strategy is employed.

The mean state of a drunk model is the expectation value of an SDE on a curved manifold. For this type of problem, it is often easiest to use *both* calculi:

In order to “get the benefit of both worlds” it is important to know how to convert an Itô equation into a Stratonovich equation, and vice versa. When calculus operations are required, conversion from Itô to the Stratonovich form can be performed, and then regular calculus can be used. Or, if expectation operations are required, a Stratonovich equation can be converted to Itô form, and then the expectation can be taken.<sup>[46]</sup>

The *Box Calculus* and *Wong-Zakai correction* can be used to convert between calculi.

### Box Calculus

The **Box Calculus** is a mnemonic for Itô's Lemma. Let  $f$  be a function of  $t, x, y, z$ . Let  $\mathbf{r}$  be a vector-valued stochastic process.<sup>3</sup> Then  $f(t, \mathbf{r}_t)$  solves the Itô SDE:

$$df(t, \mathbf{r}_t) = \dot{f}(t, \mathbf{r}_t) dt + \nabla f(t, \mathbf{r}_t) \cdot d\mathbf{r} + \frac{1}{2} d\mathbf{r}^T [\hat{D}^2 f(t, \mathbf{r}_t)] d\mathbf{r} \quad (\text{A.2})$$

<sup>3</sup>This  $\mathbf{r}$  must be a *semimartingale* and the relevant partial derivatives must exist.

Here  $\nabla f$  and  $\hat{D}^2 f$  are the gradient and Hessian second-derivative matrix of  $f$ . To remember this formula, formally Taylor-expand  $f$  to first order in  $dt$  and second order in  $dx, dy, dz$ . In column-vector notation, it is:

$$df = \partial_t f + \begin{bmatrix} \partial_x f & \partial_y f & \partial_z f \end{bmatrix} \begin{bmatrix} dx_t \\ dy_t \\ dz_t \end{bmatrix} + \frac{1}{2} \begin{bmatrix} dz_t & dy_t & dz_t \end{bmatrix} \begin{bmatrix} \partial_{xx} f & \partial_{xy} f & \partial_{xz} f \\ \partial_{yx} f & \partial_{yy} f & \partial_{yz} f \\ \partial_{zx} f & \partial_{zy} f & \partial_{zz} f \end{bmatrix} \begin{bmatrix} dx_t \\ dy_t \\ dz_t \end{bmatrix}$$

Substitute formulas for  $dx_t, dy_t, dz_t$  into (A.2) and replace any second-order infinitesimals of the form  $dt^2, dW^1 dW^1, dW^1 dt, dW^2 dW^1$ , etc. according to this box:

	$dt$	$dW^1$	$dW^2$	$dW^3$
$dt$	0	0	0	0
$dW^1$	0	$dt$	0	0
$dW^2$	0	0	$dt$	0
$dW^3$	0	0	0	$dt$

### The Wong-Zakai correction

If  $f(t, \mathbf{r}_t)$  solves (A.2), then it also solves an equivalent SF SDE:

$$df(t, \mathbf{r}_t) = \dot{f}(t, \mathbf{r}_t) dt + \nabla f(t, \mathbf{r}_t) \circ d\mathbf{r}_t$$

The SF Chain Rule is Itô's Lemma without the second-order terms. This difference can be used to derive the **Wong-Zakai correction**. Suppose  $\mathbf{r}_t$  is a solution to the SF SDE

$$d\mathbf{r}_t = \mathbf{g}(t, \mathbf{r}_t) dt + \boldsymbol{\eta}(t, \mathbf{r}_t) \circ dW$$

Let  $\hat{D}\boldsymbol{\eta}$  denote the Jacobian derivative of  $\boldsymbol{\eta}$  with respect to  $x, y, z$ . Then  $\mathbf{r}_t$  solves this Itô SDE:[113]

$$d\mathbf{r}_t = \mathbf{g}(t, \mathbf{r}_t) dt + \boldsymbol{\eta}(t, \mathbf{r}_t) dW + \frac{1}{2} [\hat{D}\boldsymbol{\eta}(t, \mathbf{r}_t)] \boldsymbol{\eta}(t, \mathbf{r}_t) dt \quad (\text{WZ})$$

This formula can be generalized to SDEs with multiple independent  $dW$  terms by summing Wong-Zakai corrections for each  $dW$ . [46] The following SDEs are equivalent:

$$d\mathbf{r}_t = \mathbf{g}(t, \mathbf{r}_t) dt + \sum_{m=1}^M \boldsymbol{\eta}_m(t, \mathbf{r}_t) \circ dW^m \quad (\text{SF})$$

$$d\mathbf{r}_t = \mathbf{g}(t, \mathbf{r}_t) dt + \sum_{m=1}^M \boldsymbol{\eta}_m(t, \mathbf{r}_t) dW^m + \frac{1}{2} \sum_{m=1}^M [\hat{D}\boldsymbol{\eta}_m(t, \mathbf{r}_t)] \boldsymbol{\eta}_m(t, \mathbf{r}_t) dt \quad (\text{It}\bar{o})$$

Here  $M$  is the number of  $dW^m$  terms and  $m$  is an index, not an exponent.

## A.2 Centrifugal drift of It $\bar{o}$ SDEs

A numerical view may provide intuition about “centrifugal drift” in It $\bar{o}$  SDEs such as

$$\begin{bmatrix} dx_t \\ dy_t \end{bmatrix} = \begin{bmatrix} 0 & \mu \\ -\mu & 0 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix} dt + \begin{bmatrix} 0 & \nu \\ -\nu & 0 \end{bmatrix} dW \quad (\text{Cyclone})$$

The lack of a  $\circ$  symbol before the  $dW$  means “use It $\bar{o}$  calculus.” If (Cyclone) were an SF SDE, the solution would be confined to a circle of radius  $\sqrt{x_0^2 + y_0^2}$ . But (Cyclone) is an It $\bar{o}$  SDE, and the expected radius grows exponentially! Consider a linear approximation with a discrete timestep of size  $h$ :

$$\begin{bmatrix} x_{t+h} \\ y_{t+h} \end{bmatrix} \approx \begin{bmatrix} x_t \\ y_t \end{bmatrix} + \begin{bmatrix} 0 & \mu \\ -\mu & 0 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix} h + \begin{bmatrix} 0 & \nu \\ -\nu & 0 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix} (W_{t+h} - W_t)$$

Let  $\Delta$  denote a random real number with standard normal distribution. Then  $\sqrt{h}\Delta$  is a random real with the same distribution as  $(W_{t+h} - W_t)$ . On each timestep, replace the Wiener increment with  $\sqrt{h}\Delta$ :

$$\begin{bmatrix} x_{t+h} \\ y_{t+h} \end{bmatrix} \approx \begin{bmatrix} 1 & \mu h + \nu \sqrt{h}\Delta \\ -(\mu h + \nu \sqrt{h}\Delta) & 1 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix}$$

This is the **Euler-Maruyama method** (EMM) from Section 5.3. Though it is neither stable nor accurate, EMM does correspond closely to It $\bar{o}$ ’s definition of  $\int dW$ . Roughly speaking, it evaluates the generator at the beginning of each timestep, just as It $\bar{o}$  intended.

The squared norm  $|\mathbf{r}_{t+h}|^2$  of the approximate solution to (Cyclone) is:

$$(1 + [\mu h + \nu \sqrt{h}\Delta]^2)(x_t^2 + y_t^2)$$

The approximate  $|\mathbf{r}_{t+h}|$  is larger than  $|\mathbf{r}_t|$ . Radial growth is a known numerical artifact of the Euler method which vanishes in the limit  $h \rightarrow 0$ . But the radial growth rate of (Cyclone) does *not* vanish as timesteps become small. Define a function  $\bar{r}^2(t) \equiv E[x_t^2 + y_t^2]$ .

The time derivative of  $\bar{r}^2(t)$  is:

$$\frac{d}{dt}\bar{r}^2(t) = \lim_{h \rightarrow 0} \frac{E[|\mathbf{r}_{t+h}|^2] - E[|\mathbf{r}_t|^2]}{h} = \lim_{h \rightarrow 0} \frac{E[(\mu h + \nu\sqrt{h}\Delta)^2]}{h}$$

The mean and variance of  $\Delta$  are  $E[\Delta] = 0$  and  $E[\Delta^2] = 1$ . The numerator is:

$$E[(\mu h + \nu\sqrt{h}\Delta)^2] = E[\mu^2 h^2] + E[2\mu\nu h\Delta] + E[\nu^2 h\Delta^2] = \mu^2 h^2 + \nu^2 h$$

The time derivative of  $\bar{r}^2(t)$  must therefore be:

$$\frac{d}{dt}\bar{r}^2(t) = \lim_{h \rightarrow 0} \frac{\mu^2 h^2 + \nu^2 h}{h} = \nu^2$$

No matter how small  $h$  is, the “centrifugal drift” remains. The solution to this ODE is:

$$\bar{r}^2(t) = \bar{r}^2(0)e^{\nu^2 t} \quad \Leftrightarrow \quad \bar{r}(t) = \bar{r}(0)e^{\frac{1}{2}\nu^2 t}$$

Suppose (Cyclone) were an SF SDE. To find an equivalent Itô SDE, add the Wong-Zakai correction:

$$\begin{bmatrix} dx_t \\ dy_t \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}\nu^2 & \mu \\ -\mu & -\frac{1}{2}\nu^2 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix} dt + \begin{bmatrix} 0 & \nu \\ -\nu & 0 \end{bmatrix} dW \quad (\text{Circle})$$

Sample paths of (Cyclone) and (Circle) are shown in Figure A.2. Both use  $\mu = 1$ ,  $\nu = 0.6$ ,  $(x_0, y_0) = (1, 0)$ ,  $t \in [0, 10]$  and the EM method with 1000 timesteps per time unit. Note that the simulation of (Circle) deviates from the unit circle by a small but visible amount. The Castell-Gaines ExpMid method is slower than EM, but it is far superior at eliminating this type of “accidental” radial drift.

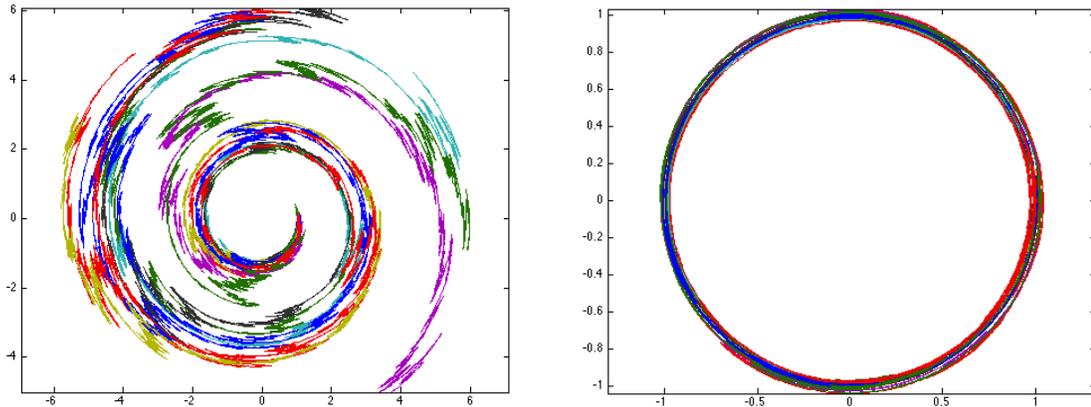


Figure A.2: Left: 10 sample paths (Cyclone). Right: 10 sample paths of (Circle).

### A.3 Probability density for spherical Brownian motion

The isotropic drunk model in Section 4.4 is an example of spherical Brownian motion. The derivation here differs considerably from Yosida's, but it leads to the same conclusion.[60] It is presented here mainly to illustrate that the two-calculi method is a much easier way to find expectation values. For the isotropic model,  $\mathbf{r}_t$  is the solution to this SF SDE:

$$d\mathbf{r}_t = \boldsymbol{\eta}_1(\mathbf{r}_t) \circ dW^1 + \boldsymbol{\eta}_2(\mathbf{r}_t) \circ dW^2 + \boldsymbol{\eta}_3(\mathbf{r}_t) \circ dW^3$$

where the superscripts  $W^1, W^2, W^3$  are indices, not exponents. The noise terms  $\boldsymbol{\eta}_m$  are:

$$\boldsymbol{\eta}_1(\mathbf{r}_t) = -\nu \hat{I} \mathbf{r}_t \quad \boldsymbol{\eta}_2(\mathbf{r}_t) = -\nu \hat{J} \mathbf{r}_t \quad \boldsymbol{\eta}_3(\mathbf{r}_t) = -\nu \hat{K} \mathbf{r}_t$$

where  $\hat{I}, \hat{J}, \hat{K}$  are the 3D rotation generators. Let  $p$  be the probability density for  $\mathbf{r}_t$ . The **Fokker-Planck-Kolmogorov (FPK) equation** is a partial differential equation for  $p$ . Formulas for the FPK equation can be found in textbooks such as [46]. For an SF SDE, and using the notation of this thesis, it is:

$$\partial_t p = \frac{1}{2} \nabla \cdot \left( \boldsymbol{\eta}_1 [\nabla \cdot (\boldsymbol{\eta}_1 p)] + \boldsymbol{\eta}_2 [\nabla \cdot (\boldsymbol{\eta}_2 p)] + \boldsymbol{\eta}_3 [\nabla \cdot (\boldsymbol{\eta}_3 p)] \right) \quad (\text{FPK-SF})$$

While it is not at all obvious in the present form, (FPK-SF) for Yosida's qubit is identical to the heat equation on a sphere. To show this, first note that the first and second radial directional derivatives of  $p$  are:

$$D_r p = \frac{\mathbf{r}}{|\mathbf{r}|} \cdot \nabla p = \frac{x \partial_x p + y \partial_y p + z \partial_z p}{\sqrt{x^2 + y^2 + z^2}}$$

$$D_r^2 p = \frac{\mathbf{r}}{|\mathbf{r}|} \cdot \nabla \left[ \frac{\mathbf{r}}{|\mathbf{r}|} \cdot \nabla p \right] = \frac{x^2 \partial_x^2 p + y^2 \partial_y^2 p + z^2 \partial_z^2 p + 2xy \partial_x \partial_y p + 2yz \partial_y \partial_z p + 2zx \partial_z \partial_x p}{x^2 + y^2 + z^2}$$

A tedious calculation (with help from *Mathematica*) shows that (FPK-SF) is:

$$\partial_t p = \frac{\nu^2}{2} \left( |\mathbf{r}|^2 \nabla^2 p - |\mathbf{r}|^2 D_r^2 p - 2|\mathbf{r}| D_r p \right)$$

In  $x, y, z$  coordinates,  $\nabla^2 p = \partial_x^2 p + \partial_y^2 p + \partial_z^2 p$ . In spherical coordinates, it is:<sup>4</sup>

$$\nabla^2 p = \partial_r^2 p + \frac{2}{r} \partial_r p + \frac{1}{r^2 \sin \theta} \partial_\theta [(\sin \theta) \partial_\theta p] + \frac{1}{r^2 (\sin \theta)^2} \partial_\phi^2 p$$

<sup>4</sup>Use the "North pole zero" conventions with  $\theta = 0$  when  $z = 1$ ,  $\phi = 0$  when  $x = 1$ , and  $\phi \in [0, 2\pi)$ .

In spherical coordinates, (FPK-SF) becomes:

$$\partial_t p = \frac{\nu^2}{2} \left( r^2 \nabla^2 p - r^2 \partial_r^2 p - 2r \partial_r p \right) = \frac{\nu^2}{2} \left( \frac{1}{\sin \theta} \partial_\theta [(\sin \theta) \partial_\theta p] + \frac{1}{(\sin \theta)^2} \partial_\phi^2 p \right) \equiv \frac{\nu^2}{2} \nabla_\Omega^2 p$$

Here  $\nabla_\Omega^2$  denotes the **Laplace-Beltrami operator** on a spherical surface – or equivalently, the  $|\mathbf{r}|^2 \nabla^2$  operator on  $\mathbb{R}^3$  minus its radial terms. The result is:

$$\partial_t p(t, \theta, \phi) = \frac{\nu^2}{2} \nabla_\Omega^2 p(t, \theta, \phi)$$

Look for separable solutions of the form  $p(t, \theta, \phi) = \mathcal{T}(t)Y(\theta, \phi)$ :

$$\frac{2\mathcal{T}'(t)}{\nu^2 \mathcal{T}(t)} = \frac{\nabla_\Omega^2 Y(\theta, \phi)}{Y(\theta, \phi)}$$

The eigenfunctions of  $\nabla_\Omega^2$  are the spherical harmonics  $Y_l^m(\theta, \phi)$  with eigenvalues  $-l(l+1)$ .

The most general solution can be written as a linear combination of eigenfunctions:

$$p(t, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l c_l^m Y_l^m(\theta, \phi) e^{-\frac{1}{2}\nu^2 l(l+1)t} \quad (\text{A.3})$$

for some constants  $\{c_l^m\}$ . Suppose the initial state is the (pure) ground state  $\theta = 0$ . The initial condition for  $p$  is a Dirac  $\delta$  distribution which can be expanded as a linear combination of spherical harmonics:[114]

$$p_0(\theta, \phi) = \delta^2(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l Y_l^m(\theta, \phi) [Y_l^m(0, 0)]^*$$

$\phi$  symmetry means only  $m=0$  terms are needed. The  $Y_l^0$  harmonics are:

$$Y_l^0(\theta) = \sqrt{\frac{2l+1}{4\pi}} P_l(\cos \theta)$$

where  $\{P_l\}$  are Legendre polynomials. The coefficients  $\{c_l^0\}$  for this initial condition are:

$$c_l^0 = [Y_l^0(0, 0)]^* = \sqrt{\frac{2l+1}{4\pi}} P_l(1) = \sqrt{\frac{2l+1}{4\pi}}$$

Substituting these coefficients into (A.3), the probability density at time  $t \geq 0$  is:

$$p(t, \theta, \phi) = \sum_{l=0}^{\infty} \frac{2l+1}{4\pi} P_l(\cos \theta) e^{-\frac{1}{2}\nu^2 l(l+1)t}$$

The mean state  $\bar{\mathbf{r}}(t)$  is the expectation value  $E[\mathbf{r}_t]$ . Axial symmetry and isotropic diffusion imply  $E[x_t] = E[y_t] = 0$ . To find  $E[z_t]$ , use  $z = r(\cos \theta)$  and  $r = 1$ , then integrate:

$$\begin{aligned} E[z_t] &= E[\cos \theta] = \int_0^\pi \int_0^{2\pi} (\cos \theta) p(t, \theta) (\sin \theta) d\phi d\theta = 2\pi \int_0^\pi (\cos \theta) p(t, \theta) (\sin \theta) d\theta \\ &= \sum_{l=0}^{\infty} \frac{2l+1}{2} e^{-\frac{1}{2}\nu^2 l(l+1)t} \int_0^\pi (\cos \theta) P_l(\cos \theta) (\sin \theta) d\theta \end{aligned}$$

The  $l = 1$  Legendre polynomial is  $P_1(\cos \theta) = \cos \theta$ . The  $\{P_l\}$  are orthogonal:

$$\int_0^\pi P_l(\cos \theta) P_{l'}(\cos \theta) (\sin \theta) d\theta = \int_{-1}^1 P_l(\cos \theta) P_{l'}(\cos \theta) d(\cos \theta) = \frac{2}{2l+1} \delta_{ll'}$$

Only the  $l = 1$  term contributes to the series for  $E[z_t]$ . The mean state is:

$$\bar{x}(t) = E[x_t] = 0 \quad \bar{y}(t) = E[y_t] = 0 \quad \bar{z}(t) = E[z_t] = e^{-\nu^2 t}$$

Rotational symmetry then implies  $E[\mathbf{r}_t]$  for any initial pure state  $\mathbf{r}_0$  must be

$$\bar{\mathbf{r}}(t) = E[\mathbf{r}_t] = \mathbf{r}_0 e^{-\nu^2 t}$$

which is the same solution predicted by the two-calculi method.

## B BASIS OBSERVABLES

Each of the models in this thesis describes a single qubit, as does the numerical code in Appendix C. This appendix suggests methods for constructing drunk models of higher-dimensional systems. The methods are based on Fano’s 1957 review [2], though many details and proofs have been omitted for brevity and simplicity.

The Pauli coordinates in Chapter 1 represent any pure or mixed qubit state using 3 real numbers and any qubit observable using 4 real numbers. For many calculations, this choice of representation leads to simpler formulas with an intuitive geometric representation. *Basis observables* can be used to define similar coordinates for  $N$ -level systems. These are  $N^2 - 1$  “favorite” observables chosen to play the role that Pauli matrices do for qubits.

### B.1 $N$ -level systems

An  $N$ -level **system** is a quantum system with  $N$  orthogonal energy eigenstates, where  $N$  is some natural number. State vectors for  $N$ -level systems are elements of the Hilbert space  $\mathbb{C}^N$ . If an ordered orthonormal basis for  $\mathbb{C}^N$  is chosen, then each state vector can be represented by an  $N \times 1$  complex column vector, and any observable can be represented by an  $N \times N$  self-adjoint complex matrix. The term *observable* will be used here to mean either an observable *or* the matrix representing that observable. This is an abuse of notation, but it is usually safe if all parties agree to use the same ordered basis for all calculations.

By this definition, a qubit is an  $N$ -level system with  $N = 2$ . A system of  $Q$  coupled qubits, also known as a **quantum register**, is an  $N$ -level system with  $N = 2^Q$ . For example, the conventional ordered basis for a system of 3 qubits is:

$$|000\rangle, |001\rangle, |010\rangle, |011\rangle, |100\rangle, |101\rangle, |110\rangle, |111\rangle$$

where e.g.  $|\Psi\rangle = |100\rangle$  means “qubit 1 is in its excited state, qubit 2 is in its ground state, and qubit 3 is in its ground state.” The system’s state vector can be any normalized superposition of these  $2^3 = 8$  basis states. Using this basis, state vectors can be represented as  $8 \times 1$  column vectors and observables as  $8 \times 8$  self-adjoint matrices.

Define the **metaspace**  $\Omega_N$  as the set of all observables of an  $N$ -level system. Given the usual rules of matrix addition and scalar multiplication,  $\Omega_N$  forms a real vector space.<sup>1</sup> The metaspace  $\Omega_N$  of an  $N$ -level system has dimension  $N^2$ . To see why, note that each matrix in  $\Omega_N$  has  $N$  real diagonal elements and  $N^2 - N$  complex off-diagonal elements. For self-adjoint matrices, each upper-triangular element is the conjugate of its corresponding lower-triangular element, so it is only necessary to specify parameters for half of the off-diagonal elements. Each complex number can be specified by two real parameters, so a total of  $N + 2 \cdot \frac{1}{2}(N^2 - N) = N^2$  independent real parameters are necessary and sufficient to uniquely specify an  $N$ -level observable.

For any two observables  $\hat{A}, \hat{B} \in \Omega_N$ , define the **Hilbert-Schmidt inner product**:<sup>2</sup>

$$\text{Tr}[\hat{A}^\dagger \hat{B}]$$

**Basis observables** are any  $N^2 - 1$  observables  $\{\hat{X}_m\} \subset \Omega_N$  which are traceless, orthogonal, and normalized such that  $\text{Tr}[\hat{X}_m^2]$  is some constant  $C$ :

$$\text{Tr}[\hat{X}_m] = 0 \qquad \text{Tr}[\hat{X}_l \hat{X}_m] = \begin{cases} C & \text{if } l = m \\ 0 & \text{if } l \neq m \end{cases}$$

Pauli matrices are an example of basis observables for  $N = 2$  with normalization  $C = 2$ . The **Gell-Mann matrices**  $\{\hat{\lambda}_m\}$  are basis observables for  $N = 3$  with  $C = 2$ :

$$\begin{aligned} \hat{\lambda}_1 &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \hat{\lambda}_2 &= \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \hat{\lambda}_3 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \hat{\lambda}_4 &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\ \hat{\lambda}_5 &= \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix} & \hat{\lambda}_6 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} & \hat{\lambda}_7 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix} & \hat{\lambda}_8 &= \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \end{aligned}$$

Basis observables can be rescaled if another normalization constant is preferred. The rest of this appendix uses the convention  $C = 1$ .

<sup>1</sup>Any *real* linear combination of  $N \times N$  self-adjoint matrices is itself an  $N \times N$  self-adjoint matrix. This is the meaning of “real vector space,” even though the elements of the space are complex matrices.

<sup>2</sup>Because the elements of  $\Omega_N$  are self-adjoint, the  $\dagger$  can be safely omitted.

## B.2 Fano coordinates

Given an  $N$ -level system, choose  $M = N^2 - 1$  basis observables  $\{\hat{X}_m\}$ . Normalize them according to the convention  $\text{Tr}[\hat{X}_m^2] = 1$ . Let  $\hat{1}$  denote the  $N \times N$  identity matrix. Define the **central observable**  $\hat{X}_0$  to be the rescaled identity matrix  $\hat{X}_0 \equiv \frac{1}{\sqrt{N}}\hat{1}$ . Then the set

$$\{\hat{X}_0, \hat{X}_1, \dots, \hat{X}_M\}$$

is an ordered orthonormal basis for  $\Omega_N$ . Any  $N$ -level observable  $\hat{A}$  can be written uniquely as a real linear combination of these  $\hat{X}$  matrices:

$$\hat{A} = a_0\hat{X}_0 + a_1\hat{X}_1 + \dots + a_M\hat{X}_M$$

$$a_0 \equiv \text{Tr}[\hat{A}\hat{X}_0] = \frac{1}{\sqrt{N}}\text{Tr}[\hat{A}] \quad a_1 \equiv \text{Tr}[\hat{A}\hat{X}_1] \quad \dots \quad a_M \equiv \text{Tr}[\hat{A}\hat{X}_M]$$

Any  $N$ -level density matrix  $\hat{\rho}$  is an  $N$ -level observable, so it can be written:

$$\hat{\rho} = r_0\hat{X}_0 + r_1\hat{X}_1 + \dots + r_M\hat{X}_M$$

The  $\hat{X}_0$  matrix has trace  $\sqrt{N}$ , and the other  $\hat{X}$  matrices are traceless. Because  $\text{Tr}[\hat{\rho}] = 1$  for any density matrix, the zeroth component must be  $r_0 = 1/\sqrt{N}$ :

$$\text{Tr}[\hat{\rho}] = r_0\text{Tr}[\hat{X}_0] = r_0\sqrt{N} = 1 \quad \Leftrightarrow \quad r_0 = \frac{1}{\sqrt{N}}$$

Define the **Fano coordinates** of  $\hat{\rho}$  to be the other  $M$  components  $\mathbf{r} \equiv (r_1, \dots, r_M)$ :

$$r_m \equiv \text{Tr}[\hat{\rho}\hat{X}_m] \quad m \in \{1, \dots, M\}$$

Any pure or mixed state can be represented by these  $M$  real numbers, each of which is the expectation value of some basis observable  $\hat{X}_m$ . Expectation values of other observables are given by dot products. For an observable  $\hat{A}$ , let  $\mathbf{A}$  denote the real vector  $(a_1, \dots, a_M)$ :

$$\langle A \rangle = \text{Tr}[\hat{\rho}\hat{A}] = r_0A_0 + r_1A_1 + \dots + r_MA_M = \frac{1}{N}\text{Tr}[\hat{A}] + \mathbf{r} \cdot \mathbf{A}$$

The condition that  $\hat{\rho}$  has non-negative eigenvalues and trace 1 requires  $\text{Tr}[\hat{\rho}^2] \leq 1$ .

$$\text{Tr}[\hat{\rho}^2] \leq 1 \quad \Leftrightarrow \quad r_0^2 + r_1^2 + \dots + r_M^2 \leq 1 \quad \Leftrightarrow \quad |\mathbf{r}| \leq 1 - \frac{1}{N}$$

Pure states have  $\text{Tr}[\hat{\rho}^2] = 1$ , which constrains these states to the hypersphere  $|\mathbf{r}|^2 = 1 - \frac{1}{N}$ . Unitary evolution requires pure states to evolve by rotations to other pure states. The maximum-entropy mixed state is  $\frac{1}{N}\hat{1}$ , which has Fano coordinates  $\mathbf{r} = \mathbf{0}$ .

Time evolution of  $\mathbf{r}$  can be found by solving  $M$  coupled Ehrenfest equations for the expectation values  $\{r_m\}$ . The Ehrenfest theorem is

$$\frac{d}{dt}\langle\hat{A}\rangle = \frac{-\imath}{\hbar}\langle[\hat{A}, \hat{H}]\rangle + \langle\partial_t\hat{A}\rangle$$

Substituting  $\hat{X}_m$  for  $\hat{A}$  and  $r_m$  for  $\langle\hat{X}_m\rangle = \text{Tr}[\hat{\rho}\hat{X}_m]$ ,

$$\hbar \frac{d}{dt}r_m = -\imath\langle[\hat{X}_m, \hat{H}]\rangle = \imath\langle[\hat{H}, \hat{X}_m]\rangle$$

The Hamiltonian  $\hat{H}$  can also be written as a linear combination of  $\hat{X}$  matrices:

$$\hat{H} = H_0\hat{X}_0 + H_1\hat{X}_1 + \cdots + H_M\hat{X}_M$$

The  $H_0$  component does not affect time evolution because every matrix commutes with the central observable  $\hat{X}_0 = \frac{1}{\sqrt{N}}\hat{1}$ . The commutator  $[\hat{H}, \hat{X}_m]$  is:

$$[\hat{H}, \hat{X}_m] = \sum_{l=1}^M H_l[\hat{X}_l, \hat{X}_m]$$

For each commutator  $[\hat{X}_l, \hat{X}_m]$ , define  $\hat{C}_{lm} \equiv \imath[\hat{X}_l, \hat{X}_m]$ . Note that each  $\hat{C}_{lm}$  matrix is traceless and self-adjoint, which means it can be uniquely specified by its Fano coordinates  $\mathbf{C}_{lm}$ . Each time evolution equation can then be written in the form:

$$\hbar \frac{d}{dt}r_m = \sum_{l=1}^M H_l\langle\hat{C}_{lm}\rangle = \sum_{l=1}^M H_l\text{Tr}[\hat{C}_{lm}\hat{\rho}] = \sum_{l=1}^M H_l\mathbf{C}_{lm} \cdot \mathbf{r}$$

Weigert pointed out that a Fano-style parameterization of  $\hat{\rho}$  and  $\hat{H}$  allows states to be represented without using state vectors or density matrices:

...the parametrization of a density matrix by expectation values suggests a conceptually interesting way to describe the time evolution of a quantum system *without* invoking its density matrix or wave function. Instead, only *expectation values* of Hermitian operators are used which can be measured directly contrary to the wave function.[12]

From a group-theory point of view,  $\{\hat{X}_m\}$  form a basis for a matrix representation of the *special unitary Lie algebra*  $\mathfrak{su}(N)$ . The *structure constants* of  $\mathfrak{su}(N)$  determine the form of the  $\hat{C}_{lm}$  matrices in the time-evolution equation. Time evolution operators are elements of the *special unitary group*  $SU(N)$  generated by  $\mathfrak{su}(N)$ . For a physics-oriented introduction to Lie groups and Lie algebras, see e.g. [115].

## C MATLAB CODE

### C.1 How to use the MATLAB code

This appendix contains the MATLAB scripts used for all simulations in Chapter 6. Each script is designed to use Pauli coordinates and dimensionless units from Chapter 1.

**StochasticLinear** and **StochasticNonlinear** are Monte Carlo simulators which use the Castell-Gaines strategy from Chapter 5 to generate sample states. The mean state  $\bar{\mathbf{r}}(t)$  is estimated by averaging all sample states. **StochasticLinear** uses the linear ExpMid method, which is faster for linear drunk models but unsuitable for nonlinear models. Nonlinear models require **StochasticNonlinear**, which uses the slower nonlinear ExpMid method. Plotting for both of these scripts is handled by **StochasticPlot**.

**StochasticPlot** script plots the mean state and one sample state produced by either **StochasticLinear** or **StochasticNonlinear**. Plots can be a time series for  $x(t), y(t), z(t)$ , trajectories on the Bloch sphere, or both. Multiple sample states can also be plotted on the Bloch sphere. This script can also plot von Neumann entropy of the mean state and, for testing purposes, radii of the first few sample states.

**DrunkenMaster** approximates the mean state  $\bar{\mathbf{r}}(t)$  of a linear drunk model. It uses the 4th-order Magnus method with Gauss-Legendre quadrature from Chapter 5 to approximately solve the master equation. It plots the mean-state  $\bar{\mathbf{r}}(t)$  in two ways: as time series for  $\bar{x}(t), \bar{y}(t), \bar{z}(t)$  and as a trajectory on the Bloch sphere.

**RamseyMaster** simulates a full Ramsey experiment with multiple waiting times. For each waiting time  $\tau_j$ , it calculates the mean state by numerically integrating the drunken master equation. Final  $\bar{z}$  coordinates are plotted versus  $\tau$ .

Near the top of each script are two function handles called **Mu** and **Sigma**. Users specify the mean field  $\boldsymbol{\mu}$  and volatility matrix  $\hat{\Sigma}$  by modifying these functions. For **StochasticLinear**, **Mu** and **Sigma** can be constants or functions of  $t$ . For **StochasticNonlinear**, **Mu** and **Sigma** can also be functions of  $x, y$ , and  $z$ . **Mu** should return a  $3 \times 1$  column of real numbers, and **Sigma** should return a  $3 \times 3$  matrix of real numbers. The scripts will automatically calculate sample times, Wong-Zakai corrections, Magnus matrices, and so on.

`InitialState` should be a  $3 \times 1$  column vector of initial Pauli coordinates  $x_0, y_0, z_0$ . Initial states are automatically normalized to have radius 1. (To use `DrunkenMaster` with a mixed-state initial condition, simply comment out the line which normalizes `InitialState`.)

Users can also specify a start time, stop time, and sample rate. As a guideline, the sample rate should be set to at least the **Nyquist rate**, which is 2 times the fastest-oscillating frequency (in Hz, not rad/sec) of the system being simulated. This ensures that matrix exponentials and Magnus series converge quickly; see Chapter 5 for details. Note that for low sample rates, the interpolation used by MATLAB's `plot` function may cause misleading plots even if the numerical results are accurate.

If  $J$  timesteps are used, then `DrunkenMaster` saves a  $1 \times (J+1)$  row called `times` and a  $3 \times (J+1)$  matrix called `mean_path`. Sample times are stored in `times`. The 1st, 2nd, and 3rd rows of `mean_path` are time series for the mean-state components  $\bar{x}, \bar{y}, \bar{z}$ . `StochasticLinear` and `StochasticNonlinear` also output a  $3 \times (J+1) \times K$  3D matrix called `all_paths`, where  $K$  is the number of sample states which were simulated. The rows  $(1, :, K)$ ,  $(2, :, K)$ ,  $(3, :, K)$  are time series for the  $x, y, z$  coordinates of the  $K$ th simulated state.

## C.2 Scripts

### StochasticLinear

```
% Simulate possible states of a drunk qubit and find their mean.
% This program is for linear models ONLY.
% Use StochasticPlot to plot the results.

%%%% User-defined parameters go here

t_start = 0;           % Initial time
t_stop = 20*pi;       % Time at which measurement is performed
sample_rate = 25;     % Number of steps per time unit
num_paths = 10;      % Number of paths to simulate

% Expected B field as 3x1 column vector. (Can be function of t.)
Mu = @(t) [ 0.1*cos(t) ; 0 ; 1 ];

% Volatility as 3x3 matrix. (Can be function of t.)
BigSigma = @(t) [ 0.05,0,0 ; 0,0.05,0 ; 0,0,0.1 ];

% Initial state as 3x1 column vector. (Will be normalized automatically.)
InitialState = [0;0;1];
InitialState = InitialState / norm(InitialState);

%%%% Simulation is here

% Make equally-spaced sample times
% Note: J steps means J+1 sample times
```

---

```

J = ceil( sample_rate * (t_stop - t_start) );
times = 0:J;
dt = 1.0 / sample_rate;
sqrt_dt = sqrt(dt);
times = t_start + dt * times;

% Define rotation generators
Imat = [0,0,0; 0,0,-1; 0,1,0];
Jmat = [0,0,1; 0,0,0; -1,0,0];
Kmat = [0,-1,0; 1,0,0; 0,0,0];

% Paths will be stored in a 3x(J+1)xK matrix.
all_paths = zeros(3,J+1,num_paths);

% For each k, simulate one possible state
disp('Simulating states...')
for k = 1:num_paths

    % Each path is a matrix. jth column is [x;y;z] at jth time.
    kth_path = zeros(3,J+1);
    kth_path(:,1) = InitialState;
    old_state = InitialState;

    % For each j, move the old state forward one step
    for j = 1:J

        % Create rescaled column of standard normal randoms
        dW = sqrt_dt * normrnd(0,1,3,1);

        % Evaluate mu and sigma at midpoint of timestep
        t_sharp = times(j) + 0.5*dt;
        new_mu = Mu(t_sharp);
        new_sigma = BigSigma(t_sharp);

        % Evolve old_state forward one step
        B = new_mu*dt + new_sigma*dW;
        G = -1.0 * ( B(1)*Imat + B(2)*Jmat + B(3)*Kmat );
        new_state = expm(G) * old_state;

        % Store new_state and recycle it for next step
        kth_path(:,j+1) = new_state;
        old_state = new_state;

    end

    % Store the new path
    all_paths(:,:,k) = kth_path;

    % Progress indicator
    if mod(k,10) == 0
        disp(k)
    end
end

end

% Mean_path is a 3x(J+1) matrix. jth column is mean state at jth time.
mean_path = mean(all_paths,3);

%% Delete temporary variables

```

```
clearvars -except times all_paths mean_path
```

## StochasticNonlinear

```
% Simulate possible states of a drunk qubit and find their mean.
% This program works for linear or nonlinear models.
% Use StochasticPlot to plot the results.

%%%%% User-defined parameters go here

t_start = 0;           % Initial time
t_stop = 20*pi;       % Time at which measurement is performed
sample_rate = 25;     % Number of steps per time unit
num_paths = 10;      % Number of paths to simulate

% Expected B field as 3x1 column vector. (Can be function of t,x,y,z.)
Mu = @(t,x,y,z) [ 0.1*cos(t) ; 0 ; 1 ];

% Volatility as 3x3 matrix. (Can be function of t,x,y,z.)
BigSigma = @(t,x,y,z) [...
    0.05*(z-0.6), 0, 0 ;...
    0, 0.05*(z-0.6), 0 ;...
    0, 0, 0.1 ];

% Initial state as 3x1 column vector. (Will be normalized automatically.)
InitialState = [0;0;1];
InitialState = InitialState / norm(InitialState);

%%%%% Simulate qubit paths

% Make equally-spaced sample times
% Note: J steps means J+1 sample times
J = ceil( sample_rate * (t_stop - t_start) );
times = 0:J;
dt = 1.0 / sample_rate;
sqrt_dt = sqrt(dt);
times = t_start + dt * times;

% Define rotation generators
Imat = [0,0,0; 0,0,-1; 0,1,0];
Jmat = [0,0,1; 0,0,0; -1,0,0];
Kmat = [0,-1,0; 1,0,0; 0,0,0];

% Paths will be stored in a 3x(J+1)xK matrix.
all_paths = zeros(3,J+1,num_paths);

% For each k, simulate one complete sample path
disp('Simulating states...')
for k = 1:num_paths

    % Each path is a matrix. jth column is [x;y;z] at jth time.
    kth_path = zeros(3,J+1);
    kth_path(:,1) = InitialState;
    old_state = InitialState;

    % For each j, generate an increment of B and calculate a new state
```

```

for j = 1:J

    % Create rescaled column of standard normal randoms
    dW = sqrt_dt * normrnd(0,1,3,1);

    % Evaluate mu and sigma at left end of timestep
    t_old = times(j);
    x_old = old_state(1);
    y_old = old_state(2);
    z_old = old_state(3);
    mu_old = Mu(t_old,x_old,y_old,z_old);
    sigma_old = BigSigma(t_old,x_old,y_old,z_old);
    B_old = mu_old*dt + sigma_old*dW;

    % Find midpoint state
    G_old = -0.5 * ( B_old(1)*Imat + B_old(2)*Jmat + B_old(3)*Kmat );
    r_mid = expm(G_old) * old_state;

    % Evaluate mu and sigma at midpoint of timestep
    t_mid = t_old + 0.5*dt;
    x_mid = r_mid(1);
    y_mid = r_mid(2);
    z_mid = r_mid(3);
    mu_mid = Mu(t_mid,x_mid,y_mid,z_mid);
    sigma_mid = BigSigma(t_mid,x_mid,y_mid,z_mid);
    B_mid = mu_mid*dt + sigma_mid*dW;

    % Evolve old_state forward one step
    G_mid = -1.0*(B_mid(1)*Imat + B_mid(2)*Jmat + B_mid(3)*Kmat);
    new_state = expm(G_mid) * old_state;

    % Store new_state and recycle it for next step
    kth_path(:,j+1) = new_state;
    old_state = new_state;

end

% Store the new path
all_paths(:,j,k) = kth_path;

% Progress indicator
if mod(k,10) == 0
    disp(k)
end

end

% Mean_path is a 3x(J+1) matrix. jth column is mean state at jth time.
mean_path = mean(all_paths,3);

%%%% Delete temporary variables
clearvars -except times all_paths mean_path

```

## StochasticPlot

```

% Plot output of DrunkLinear or DrunkNonlinear solvers.

```

```

%%%%% User-defined parameters go here

plot_time_series = true;    % Show time series plot?

plot_Bloch_ball = true;    % Show Bloch-sphere trajectories?
num_sample_states = 1;     % How many sample states to draw on sphere?

plot_entropy = true;      % Show von Neumann entropy of mean state?

plot_radius = true;       % Used to check if pure states stay pure
num_radial_states = 1;    % How many states' radii to show?

close all

%%%%% Time series plot of mean state and a sample state
if plot_time_series
    t_start = times(1);
    t_stop = times(end);
    f1 = figure('Position',[0,700,1000,250]);
    set(gcf,'color',[1 1 1])

    % Plot first sample state
    plot(...
        times,all_paths(1,:,1),'b',...
        times,all_paths(2,:,1),'g',...
        times,all_paths(3,:,1),'r')
    axis([t_start,t_stop,-1,1])
    hold all

    % Plot mean state
    plot(...
        times,mean_path(1,:), 'b',...
        times,mean_path(2,:), 'g',...
        times,mean_path(3,:), 'r',...
        'LineWidth',1.5)
    hold off
end

%%%%% Bloch-ball plots of mean state and sample states
if plot_Bloch_ball
    % Plot mean state in black
    f2 = figure('Position',[1000,550,400,400]);
    set(gcf,'color',[1 1 1])
    plot3(...
        mean_path(1,:),...
        mean_path(2,:),...
        mean_path(3,:),...
        'LineWidth',1,'Color','k')
    axis equal tight
    axis([-1,1,-1,1,-1,1])

    % Plot sample states
    f3 = figure('Position',[1000,0,400,400]);
    set(gcf,'color',[1 1 1])
    axis equal tight
    axis([-1,1,-1,1,-1,1])
    hold on

```

```

    for k = 1:num_sample_states
        plot3(...
            all_paths(1,:,k),...
            all_paths(2,:,k),...
            all_paths(3,:,k),...
            'LineStyle','-', 'Color','b');
    end
    hold off
end

%%%% Plot von Neumann entropy of mean state
if plot_entropy
    f4 = figure('Position',[0,350,1000,250]);
    set(gcf,'color',[1 1 1])
    binary_entropy = @(p) -p.*log2(p) - (1-p).*log2(1-p);
    mean_r_squared = ...
        mean_path(1,:).^2 ...
        + mean_path(2,:).^2 ...
        + mean_path(3,:).^2 ;
    mean_r = sqrt(mean_r_squared);
    p_series = 0.5*(1+mean_r);
    vN_entropy = binary_entropy(p_series);
    plot(times,real(vN_entropy),'k','LineWidth',1.5)
    axis([t_start,t_stop,0,1])
end

%%%% Plot radial error of sample states
if plot_radius
    f5 = figure('Position',[0,0,1000,250]);
    set(gcf,'color',[1 1 1])
    for path = 1:num_radial_states
        r_squared = ...
            all_paths(1,:,path).^2 ...
            + all_paths(2,:,path).^2 ...
            + all_paths(3,:,path).^2 ;
        radial_error = sqrt(r_squared) - 1.0;
        plot(times,radial_error)
        hold all
    end
    hold off
end

%%%% Delete temporary variables
clearvars -except times all_paths mean_path vN_entropy

```

## DrunkenMaster

```

% Solve the master equation for the mean state of a drunk model.
% This program is for linear models ONLY.

%%%% User-defined parameters go here

t_start = 0;           % Initial time
t_stop = 40*pi;       % Time at which measurement is performed

```

```

sample_rate = 10;    % Number of steps per time unit

% Expected B field as 3x1 column vector. (Can be function of t.)
Mu = @(t) [ 0.1*cos(t) ; 0 ; 1 ];

% Volatility as 3x3 matrix. (Can be function of t.)
Sigma = @(t) [ 0.05,0,0 ; 0,0.05,0 ; 0,0,0.1 ];

% Initial state as 3x1 column vector. (Will be normalized automatically.)
InitialState = [0;0;1];
InitialState = InitialState / norm(InitialState);

%%%% Simulation is here

% Make equally-spaced sample times
% Note: J steps means J+1 sample times
J = ceil( sample_rate * (t_stop - t_start) );
times = 0:J;
dt = 1.0 / sample_rate;
times = t_start + dt * times;

% mean_path is a matrix. jth column is mean [x;y;z] at jth time.
mean_path = zeros(3,J+1);
mean_path(:,1) = InitialState;
old_state = InitialState;

% Define rotation generators and basis vectors
Imat = [0,0,0; 0,0,-1; 0,1,0];
Jmat = [0,0,1; 0,0,0; -1,0,0];
Kmat = [0,-1,0; 1,0,0; 0,0,0];

% This function will use Mu and BigSigma to find the generator
G = @(mu,sig) -1.0*( mu(1)*Imat + mu(2)*Jmat + mu(3)*Kmat )...
    +0.5*( sig'*sig - norm(sig,'fro')^2*eye(3) );

% Calculate constants for Magnus method
c_left = 0.5 - sqrt(3)/6.0;
c_right = 0.5 + sqrt(3)/6.0;
cfactor = -1.0*sqrt(3)/12;

% For each j, move the old mean state forward one step
for j = 1:J

    % Evaluate Mu and BigSigma at subsample times
    t_L = times(j) + c_left*dt;
    mu_L = Mu(t_L);
    sig_L = Sigma(t_L);
    G_L = G(mu_L,sig_L);

    t_R = times(j) + c_right*dt;
    mu_R = Mu(t_R);
    sig_R = Sigma(t_R);
    G_R = G(mu_R,sig_R);

    % Calculate Magnus matrix
    C = G_L*G_R - G_R*G_L;
    Omega = 0.5*dt*(G_L + G_R) + dt*dt*cfactor*C;

```

```

    % Evolve old_state forward one step
    new_state = expm(Omega) * old_state;

    % Store new_state and recycle it for next step
    mean_path(:,j+1) = new_state;
    old_state = new_state;

end

%%%%% Plotting stuff

close all

% Plot mean state as time series
f1 = figure(1);
set(f1, 'Position', [500,600,800,200]);
set(gcf, 'color', [1 1 1])
axis([t_start, t_stop, -1, 1])
hold on
plot(times, mean_path(1,:), 'b');
plot(times, mean_path(2,:), 'g');
plot(times, mean_path(3,:), 'r');
legend('x', 'y', 'z');
hold off

% Plot mean state in Bloch ball
f2 = figure(2);
set(f2, 'Position', [500,0,500,500]);
set(gcf, 'color', [1 1 1])
plot3(...
    mean_path(1,:),...
    mean_path(2,:),...
    mean_path(3,:),...
    'Color', 'k')
axis equal tight
axis([-1,1,-1,1,-1,1])

%%%%% Delete temporary variables
clearvars -except mean_path times

```

## RamseyMaster

```

% Simulate a Ramsey-fringe experiment using the master equation.
% This program is for linear models ONLY.

%%%%% User-defined parameters go here

max_tau = 4*pi;           % Longest waiting time to use
omega0 = 1;               % Qubit's natural angular frequency
amp = 0.05;               % Pulse amplitude
detuning = 0;             % Pulse detuning
sample_rate = 10;        % Number of steps per time unit
InitialState = [0;0;1];   % Initial state (not automatically normalized)

% Volatility as 3x3 matrix. (Can be function of t.)

```

```

Sigma = @(t) [ 0.1,0,0 ; 0,0.1,0 ; 0,0,0.2 ];

%%%%% Calculate useful constants and define anonymous functions

% Useful numbers to calculate
% Note: num_taus is arbitrarily set to 12 for nice-looking plots.
omega1 = omega0 + detuning;
pulse_t = 0.5*pi/amp;
full_cycle = 2.0*pi/omega0;
num_taus = ceil(12.0*max_tau/full_cycle);

% Decide which values of tau to use and store them as a column
tau_list = linspace(0,max_tau,num_taus)';

% Store the final z coordinates as a column
final_z = zeros(num_taus,1);

% Rotation generators
Imat = [0,0,0; 0,0,-1; 0,1,0];
Jmat = [0,0,1; 0,0,0; -1,0,0];
Kmat = [0,-1,0; 1,0,0; 0,0,0];

% Constants for Magnus method
dt = 1.0 / sample_rate;
c_left = 0.5 - sqrt(3)/6.0;
c_right = 0.5 + sqrt(3)/6.0;
cfactor = -1.0*sqrt(3)/12;

% Mean field is a piecewise function
Mu = @(t,tau) (t<=pulse_t)*[ 2.0*amp*cos(omega1*t) ; 0 ; omega0 ]...
    +(t>pulse_t)*(t<=pulse_t+tau)*[ 0 ; 0 ; omega0]...
    +(t>pulse_t+tau)*...
    [2.0*amp*cos(omega1*(t-tau-pulse_t));0;omega0];

% This function will use Mu and BigSigma to find the generator
G = @(mu,sig) -1.0*( mu(1)*Imat + mu(2)*Jmat + mu(3)*Kmat )...
    +0.5*( sig'*sig - norm(sig,'fro')^2*eye(3) );

%%%%% Simulation is here

disp('Calculating');
disp(num_taus);
disp('Ramsey fringes...');

% For each k, solve the master equation for a particular tau
% (This is inefficient, but often fast enough.)
for k = 1:num_taus

    % Make equally-spaced sample times
    t_stop = tau_list(k) + 2.0*pulse_t;
    J = ceil( sample_rate * t_stop );
    times = 0:J;
    times = dt * times;

    % Reset old_state and update tau
    old_state = InitialState;
    new_tau = tau_list(k);

```

---

```

% For each j, move the old mean state forward one step
for j = 1:J

    % Evaluate Mu and BigSigma at subsample times
    t_L = times(j) + c_left*dt;
    mu_L = Mu(t_L,new_tau);
    sig_L = Sigma(t_L);
    G_L = G(mu_L,sig_L);

    t_R = times(j) + c_right*dt;
    mu_R = Mu(t_R,new_tau);
    sig_R = Sigma(t_R);
    G_R = G(mu_R,sig_R);

    % Calculate Magnus matrix
    C = G_L*G_R - G_R*G_L;
    Omega = 0.5*dt*(G_L + G_R) + dt*dt*cfactor*C;

    % Evolve old_state forward one step
    new_state = expm(Omega) * old_state;

    % Store new_state and recycle it for next step
    old_state = new_state;

end

% Save the final z coordinate
final_z(k) = new_state(3);

% Progress indicator
if mod(k,10) == 0
    disp(k)
end

end

%%%% Plotting stuff

close all

% Plot final_z versus tau
f1 = figure(1);
set(f1,'Position',[500,600,800,200]);
set(gcf,'color',[1 1 1])
axis([tau_list(1),tau_list(end),-1,1])
hold on
plot(tau_list,final_z,'k.-');
legend('Final z','Location','SouthEast')
hold off

%%%% Delete temporary variables
clearvars -except mean_path times

```

## LIST OF REFERENCES

- [1] J. von Neumann (translated from German by R. T. Beyer), *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, 1955. [0.1](#), [1.1](#), [1.3](#), [2](#), [2.4](#)
- [2] U. Fano, “Description of states in quantum mechanics by density matrix and operator techniques,” *Rev. Mod. Phys.*, vol. 29, no. 1, 1957. [1](#), [1.1](#), [1.2](#), [1.2](#), [2.4](#), [7.3](#), [B](#)
- [3] A. J. Leggett, “Superconducting qubits – a major roadblock dissolved?,” *Science*, vol. 296, May 2002. [1.2](#), [6](#)
- [4] M. D. Shuster, “A survey of attitude representations,” *J. of the Astronautical Sciences*, vol. 41, no. 4, p. 496, 1993. [1.3](#)
- [5] D. Leibfried, R. Blatt, C. Monroe, and D. Wineland, “Quantum dynamics of single trapped ions,” *Rev. Mod. Phys.*, vol. 75, 2003. [1.4](#), [1](#)
- [6] J. K. Asbóth, P. Adam, M. Koniorczyk, and J. Janszky, “Coherent-state qubits: entanglement and decoherence,” *Eur. Phys. J. D*, vol. 30, pp. 403–410, 2004. [1.4](#)
- [7] M. H. Devoret, A. Wallraff, and J. M. Martinis, “Superconducting qubits: A short review,” *arXiv:cond-mat/0411174*, Nov 2004. [1.4](#)
- [8] M. H. Devoret and R. J. Schoelkopf, “Superconducting circuits for quantum information: An outlook,” *Science*, vol. 339, Mar 2013. [1.4](#)
- [9] A. Lengwenus, J. Kruse, M. Volk, W. Ertmer, and G. Birkl, “Coherent manipulation of atomic qubits in optical micropotentials,” *Appl. Phys. B*, vol. 86, 2007. [1.4](#), [6.4](#), [6.12](#)
- [10] P. C. Maurer, G. Kucsko, C. Latta, L. Jiang, N. Y. Yao, S. D. Bennett, F. Pastawski, D. Hunger, N. Chisholm, M. Markham, D. J. Twitchen, J. I. Cirac, and M. D. Lukin, “Room-temperature quantum bit memory exceeding one second,” *Science*, vol. 336, Jun 2012. [1.4](#)
- [11] Y. Aharonov, D. Z. Albert, and L. Vaidman, “How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100,” *Phys. Rev. Lett.*, vol. 60, no. 14, pp. 1351–1354, 1988. [2](#)
- [12] S. Weigert, “Quantum time evolution in terms of nonredundant probabilities,” *Phys. Rev. Lett.*, vol. 84, no. 5, 2000. [3](#), [7.3](#), [B.2](#)
- [13] É. Borel, “Les probabilités dénombrables et leurs applications arithmétiques,” *Rendiconti del Circolo Matematico di Palermo*, vol. 27, no. 1, pp. 247–271, 1909. [2.2](#)
- [14] C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal*, vol. 27, 1948. [2.4](#), [2.4](#)

- 
- [15] M. Tribus and E. C. McIrvine, “Energy and information,” *Scientific American*, vol. 225, no. 3, 1971. [2.4](#)
- [16] I. I. Hirschman, Jr., “A note on entropy,” *Am. J. Math.*, vol. 79, no. 1, pp. 152–156, 1957. [2.4](#)
- [17] H. Everett III, *The Theory of the Universal Wave Function*. PhD thesis, Princeton University, 1956. [2.4](#)
- [18] W. Beckner, “Inequalities in Fourier analysis,” *Annals of Mathematics, 2nd series*, vol. 102, no. 1, pp. 159–182, 1975. [2.4](#)
- [19] S. Wehner and A. Winter, “Entropic uncertainty relations – a survey,” *New J. Phys.*, vol. 12, no. 2, 2010. [2.4](#)
- [20] W. H. Zurek, “Decoherence and the transition from quantum to classical - revisited,” *Los Alamos Science*, 2002. [2.4](#), [1](#)
- [21] M. Plischke and B. Bergersen, *Equilibrium Statistical Physics*, ch. 2, pp. 46–51. World Scientific, 2nd ed., 1999. [2.5](#)
- [22] D. V. Schroeder, *An Introduction to Thermal Physics*, ch. 3. Addison Wesley Longman, 2000. [2.5.1](#)
- [23] S. Carroll, *From Eternity to Here: The Quest for the Ultimate Theory of Time*. Penguin Group, New York, 2010. [1](#)
- [24] E. T. Jaynes, “Gibbs vs Boltzmann entropies,” *Am. J. Phys.*, vol. 33, no. 391, 1965. [3](#), [7.1.1](#)
- [25] S. Kennerly, “Illusory decoherence,” *Foundations of Physics*, vol. 42, September 2012. [3](#)
- [26] C. J. Myatt, B. E. King, Q. A. Turchette, C. A. Sackett, D. Kielpinski, W. M. Itano, C. Monroe, and D. J. Wineland, “Decoherence of quantum superpositions through coupling to engineered reservoirs,” *Nature*, vol. 403, no. 20, 2000. [3](#), [7.1.2](#), [7.1.3](#)
- [27] G. van der Zouw, M. Weber, J. Felber, R. Gähler, P. Geltenbort, and A. Zeilinger, “Aharonov-Bohm and gravity experiments with the very-cold-neutron interferometer,” *Nuclear Instruments and Methods in Physics Research A*, vol. 440, 2000. [3.1](#), [4](#)
- [28] J. Wheeler and W. Zurek, eds., *Quantum Theory and Measurement*. Princeton University Press, 1983. [3](#)
- [29] V. Jacques, E. Wu, F. Grosshans, F. Treussart, P. Grangier, A. Aspect, and J.-F. Roch, “Experimental realization of Wheeler’s delayed-choice gedankenexperiment,” *Science*, vol. 315, February 2007. [3](#)
- [30] M. Chapman, C. Ekstrom, T. Hammond, R. Rubenstein, J. Schmiedmayer, S. Wehinger, and D. Pritchard, “Optics and interferometry with Na<sub>2</sub> molecules,” *Phys. Rev. Lett.*, vol. 74, no. 24, 1995. [3.1](#)
- [31] O. Nairz, M. Arndt, and A. Zeilinger, “Quantum interference experiments with large molecules,” *Am. J. Phys.*, vol. 71, no. 4, 2003. [3.1](#)

- 
- [32] S. Gerlich, S. Eibenberger, M. Tomandl, S. Nimmrichter, K. Hornberger, P. J. Fagan, J. Tüxen, M. Mayor, and M. Arndt, “Quantum interference of large organic molecules,” *Nature Communications*, vol. 2, no. 263, 2011. [3.1](#), [4](#)
- [33] W. D. Oliver, Y. Yu, J. C. Lee, K. K. Berggren, L. S. Levitov, and T. P. Orlando, “Mach-Zehnder interferometry in a strongly driven superconducting qubit,” *Science*, pp. 1653–1657, December 2005. [3.1](#)
- [34] E. Collin, G. Ithier, A. Aassime, P. Joyez, D. Vidon, and D. Esteve, “NMR-like control of a quantum bit superconducting circuit,” *Phys. Rev. Lett.*, vol. 93, no. 15, 2004. [3.1](#)
- [35] A. Zeilinger, “General properties of lossless beam splitters in interferometry,” *American Journal of Physics*, vol. 49, no. 9, 1981. [3.1.1](#)
- [36] C. E. Shannon, “Communication theory of secrecy systems,” *Bell System Technical Journal*, vol. 28, no. 4, pp. 656–715, 1949. [5](#)
- [37] S. N. Molokov, “Quantum cryptography and V. A. Kotel’nikov’s one-time key and sampling theorems,” *Physics-Uspeski*, vol. 49, no. 7, pp. 750–761, 2006. [5](#)
- [38] A. Miffre, M. Jacquy, M. Büchner, G. Tréneç, and J. Vigué, “Vibration induced phase noise in Mach-Zehnder atom interferometers,” *Appl. Phys. B*, vol. 84, pp. 617–625, 2006. [3.2](#)
- [39] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, “Experimental realization of any discrete unitary operator,” *Phys. Rev. Lett.*, vol. 73, no. 1, 1994. [3.2.2](#)
- [40] A. Abragam, *The Principles of Nuclear Magnetism*, ch. III.A. Oxford University Press, 1961. [3.3](#)
- [41] E. P. Wigner, “Characteristic vectors of bordered matrices with infinite dimensions,” *Annals of Mathematics*, vol. 62, no. 3, 1955. [4](#)
- [42] Y. Tanimura and R. Kubo, “Time evolution of a quantum system in contact with a nearly Gaussian-Markoffian noise bath,” *J. Phys. Soc. Japan*, vol. 58, no. 1, pp. 101–114, 1988. [4](#), [4](#)
- [43] N. G. van Kampen, “Itô versus Stratonovich,” *J. Stat. Phys.*, vol. 24, no. 1, 1981. [4](#), [4.1](#)
- [44] H. J. Sussmann, “On the gap between deterministic and stochastic ordinary differential equations,” *The Annals of Probability*, vol. 6, no. 1, 1978. [4](#)
- [45] H. J. Sussmann, “Limits of the Wong-Zakai type with a modified drift term,” *Stochastic Analysis*, pp. 475–493, 1991. [4](#)
- [46] G. S. Chirikjian, *Stochastic Models, Information Theory, and Lie Groups*, ch. 8: Stochastic Processes on Manifolds. Birkhäuser, 2009. [1](#), [7.3](#), [A](#), [A.1](#), [A.1](#), [A.3](#)
- [47] M. H. Devoret and J. M. Martinis, “Implementing qubits with superconducting integrated circuits,” *Quantum Information Processing*, vol. 3, no. 1-5, 2004. [4](#), [4.5](#), [1](#)
- [48] R. Kubo and N. Hashitsume, “Brownian motion of spins,” *Progress of Theoretical Physics (Supplement)*, vol. 46, pp. 210–220, 1970. [4](#)

- 
- [49] R. F. Fox, “Gaussian stochastic processes in physics,” *Physics Reports*, vol. 48, no. 3, pp. 179–283, 1978. [4](#), [7.3](#)
- [50] R. F. Fox, “Physical applications of multiplicative stochastic processes II: derivation of the Bloch equations for magnetic relaxation,” *J. Math. Phys.*, vol. 14, no. 2, 1974. [4](#), [7.3](#)
- [51] R. F. Fox, “An “H-theorem” for multiplicative stochastic processes,” *J. Math. Phys.*, vol. 13, no. 11, 1972. [4](#), [7.3](#)
- [52] V. Gorini and A. Kossakowski, “Completely positive dynamical semigroups of N-level systems,” *J. Math. Phys.*, vol. 17, no. 5, 1976. [4](#), [7.2.1](#)
- [53] V. Gorini and A. Kossakowski, “N-level system in contact with a singular reservoir,” *J. Math. Phys.*, vol. 17, no. 7, 1976. [4](#), [7.3](#)
- [54] J. T. Stockburger, “Simulating spin-boson dynamics with stochastic Liouville-von Neumann equations,” *Chemical Physics*, vol. 296, pp. 159–169, 2004. [4](#)
- [55] J. T. Stockburger and C. H. Mak, “A stochastic Liouvillian algorithm to simulate dissipative quantum dynamics with arbitrary precision,” *arXiv:cond-mat/9811277v1 [cond-mat.stat-mech]*, 1998. [4](#)
- [56] Y. Tanimura, “Stochastic Liouville, Langevin, Fokker-Planck, and master equation approaches to quantum dissipative systems,” *J. Phys. Soc. Japan*, vol. 75, no. 8, 2006. [4](#)
- [57] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*. Springer, 6 ed., 2010. [3](#), [4.6](#), [7.3](#), [A](#), [2](#)
- [58] J. M. Steele, *Stochastic Calculus and Financial Applications*. Springer, 2001. [3](#), [4.6](#), [7.3](#), [A](#), [2](#)
- [59] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, ch. 4: Stochastic Differential Equations. Springer-Verlag, 1992. [3](#), [4](#), [7.3](#), [A](#)
- [60] K. Yosida, “Brownian motion on the surface of the 3-sphere,” *Annals of Math. Stat.*, vol. 20, no. 2, pp. 292–296, 1949. [4.4](#), [A.3](#)
- [61] J. Clarke and F. K. Wilhelm, “Superconducting quantum bits,” *Nature*, vol. 453, June 2008. [4.5](#)
- [62] G. Wendin and V. S. Shumeiko, “Superconducting quantum circuits, qubits and computing,” *arXiv:cond-mat/0508729v1*, 2005. [4.5](#), [1](#)
- [63] J. C. Hull, *Options, Futures, and Other Derivatives*, ch. 14. Prentice Hall, 8th ed., 2012. [10](#)
- [64] W. Magnus, “On the exponential solution of differential equations for a linear operator,” *Comm. Pure and Appl. Math*, vol. 4, pp. 649–673, 1954. [5.1](#)
- [65] E. B. Dynkin, “Calculation of the coefficients in the Campbell-Hausdorff formula,” *Dokl. Akad. Nauk SSSR (N.S.)*, vol. 57, pp. 323–326, 1947. [5.1](#), [5.1.1](#)
- [66] A. Iserles, “Magnus expansions and beyond,” *Contemporary Maths*, vol. 539, 2011. [5.1](#), [5.1.2](#)

- 
- [67] M. Hochbruck and C. Lubich, “On Magnus integrators for time-dependent Schrödinger equations,” *SIAM J. Numerical Analysis*, vol. 41, no. 3, 2004. [5.1](#), [2](#), [4](#)
- [68] M. Khanamiryan, *Numerical methods for systems of highly oscillatory ordinary differential equations*. PhD thesis, U. of Cambridge, (2010). [2](#)
- [69] S. Olver, *Numerical approximation of highly oscillatory integrals*. PhD thesis, U. of Cambridge, 2008. [2](#)
- [70] M. Hochbruck and A. Ostermann, “Exponential integrators,” *Acta Numerica*, 2010. [5.1](#), [4](#)
- [71] R. P. Feynman, “Simulating physics with computers,” *Intl. J. of Theoretical Physics*, vol. 21, no. 6/7, 1982. [5.1](#)
- [72] H. F. Baker, “Alternants and continuous groups,” *Proc. London Math. Soc.*, vol. 2, no. 3, p. 24, 1905. [5.1.1](#)
- [73] J. E. Campbell, “On a law of combination of operators bearing on the theory of continuous transformation groups,” *Proc. London Math. Soc.*, vol. 1, no. 28, p. 381, 1897. [5.1.1](#)
- [74] F. Hausdorff *Leipziger Berichte*, vol. 58, 1906. [5.1.1](#)
- [75] R. Gilmore, “Baker-Campbell-Hausdorff formulas,” *J. Math. Phys.*, vol. 15, no. 12, 1974. [3](#)
- [76] S. Blanes, F. Casas, J. Oteo, and J. Ros, “The Magnus expansion and some of its applications,” *Physics Reports*, vol. 470, 2009. [5.1.1](#), [5.1.2](#)
- [77] A. Iserles, H. Munthe-Kaas, S. Nørsett, and A. Zanna, “Lie-group methods,” *Acta Numerica*, pp. 215–365, 2000. [5.1.1](#), [5.1.2](#)
- [78] C. Moler and C. van Loan, “Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later.,” *SIAM Review*, vol. 45, no. 1, 2003. [5.2](#), [5.2](#)
- [79] K. Burrage and P. M. Burrage, “High strong order methods for non-commutative stochastic ordinary differential equation systems and the magnus formula,” *Physica D*, vol. 133, pp. 34–48, 1999. [5.3](#)
- [80] J. M. C. Clark and R. J. Cameron, *Stochastic Differential Systems Filtering and Control*, vol. 25 of *Lecture Notes in Control and Information Sciences*, ch. The maximum rate of convergence of discrete approximations for stochastic differential equations. Springer-Verlag, 1980. [5.3](#)
- [81] F. Castell and J. Gaines, “The ordinary differential equation approach to asymptotically efficient schemes for solution of stochastic differential equations,” *Annales de l’I. H. P., section B*, vol. 32, no. 2, pp. 231–250, 1996. [5.3.1](#)
- [82] I. I. Rabi, “Space quantization in a gyrating magnetic field,” *Phys. Rev.*, vol. 51, pp. 652–654, April 1937. [6](#)
- [83] N. F. Ramsey, “A molecular beam resonance method with separated oscillating fields,” *Phys. Rev.*, vol. 78, no. 6, 1950. [6](#)

- 
- [84] M. Brune, F. Schmidt-Kaler, A. Maali, J. Dreyer, E. Hagley, J. M. Raimond, and S. Haroche, “Quantum Rabi oscillation: A direct test of field quantization in a cavity,” *Phys. Rev. Lett.*, vol. 76, no. 11, 1996. [6](#)
- [85] D. M. Meekhof, C. Monroe, B. E. King, W. M. Itano, and D. J. Wineland, “Generation of nonclassical motional states of a trapped atom,” *Phys. Rev. Lett.*, vol. 76, no. 1, 1996. [6](#)
- [86] D. J. Griffiths, *Introduction to Quantum Mechanics*. Prentice Hall, 1995. [6.1.1](#)
- [87] E. L. Hahn, “Free nuclear induction,” *Physics Today*, November 1953. [6.1.2](#), [6.1](#), [3](#)
- [88] J. J. Pla, K. Y. Tan, J. P. Dehollain, W. H. Lim, J. J. L. Morton, F. A. Zwanenburg, D. N. Jamieson, A. S. Dzurak, and A. Morello, “High-fidelity readout and control of a nuclear spin qubit in silicon,” *Nature*, vol. 496, April 2013. [6.4](#), [6.10](#)
- [89] H. Paik, S. K. Dutta, R. M. Lewis, T. A. Palomake, B. K. Cooper, R. C. Ramos, H. Xu, A. J. Dragt, J. R. Anderson, C. J. Lobb, and F. C. Wellstood, “Decoherence in dc SQUID phase qubits,” *Phys. Rev. B*, vol. 77, 2008. [6.4](#), [6.11](#)
- [90] I. Chiorescu, Y. Nakamura, C. J. P. M. Harmans, and J. E. Mooij, “Coherent quantum dynamics of a superconducting flux qubit,” *Science*, vol. 299, Mar 2003. [6.4](#), [6.13](#)
- [91] D. D. Awschalom, L. C. Bassett, A. S. Dzurak, E. L. Hu, and J. R. Petta, “Quantum spintronics: Engineering and manipulating atom-like spins in semiconductors,” *Science*, vol. 339, Mar 2013. [1](#)
- [92] J. M. Raimond, M. Brune, and S. Haroche, “Manipulating quantum entanglement with atoms and photons in a cavity,” *Rev. Mod. Phys.*, vol. 73, 2001. [1](#)
- [93] A. S. Eddington, *The Nature of the Physical World*. The University Press, 1929. [7.1](#), [7.1.2](#)
- [94] E. T. Jaynes, “Information theory and statistical mechanics,” *Phys. Rev.*, vol. 106, May 1957. [7.1.1](#)
- [95] E. T. Jaynes, “Information theory and statistical mechanics II,” *The Physical Review*, vol. 108, no. 2, pp. 171–190, 1957. [7.1.1](#), [7.3](#)
- [96] P. Bocchieri and A. Loinger, “Quantum recurrence theorem,” *Phys. Rev.*, vol. 107, 1957. [7.1.2](#)
- [97] I. C. Percival, “Almost periodicity and the quantal H theorem,” *J. Math. Phys.*, vol. 2, 1961. [7.1.2](#)
- [98] L. S. Schulman, “Note on the quantum recurrence theorem,” *Phys. Rev. A*, vol. 18, 1978. [7.1.2](#)
- [99] E. Joos, H. D. Zeh, C. Kiefer, D. Giulini, J. Kupsch, and I.-O. Stamatescu, *Decoherence and the Appearance of a Classical World in Quantum Theory*. Springer, 2 ed., 2003. [7.1.3](#), [7.1.3](#)
- [100] A. M. Jayannavar, “Comment on some theories of state reduction,” *Physics Letters A*, vol. 167, 1992. [7.1.3](#)

- 
- [101] Q. A. Turchette, C. J. Myatt, B. E. King, C. A. Sackett, D. Kielpinski, W. M. Itano, C. Monroe, and D. J. Wineland, “Decoherence and decay of motional quantum states of a trapped atom coupled to engineered reservoirs,” *Phys. Rev. A*, vol. 62, 2000. [7.1.3](#)
- [102] C. Monroe, D. M. Meekhof, B. E. King, and D. J. Wineland, “A ‘Schrödinger cat’ superposition state of an atom,” *Science*, vol. 272, 1996. [7.1.3](#)
- [103] W. H. Zurek, “Decoherence, einselection, and the quantum origins of the classical,” *Rev. Mod. Phys.*, vol. 75, Jul 2003. [7.1.3](#), [7.2](#)
- [104] R. P. Feynman and F. L. Vernon Jr., “The theory of a general quantum system interacting with a linear dissipative system,” *Annals of Physics*, vol. 24, pp. 118–173, 1963. [7.2](#)
- [105] A. O. Caldeira and A. J. Leggett, “Influence of damping on quantum interference: An exactly soluble model,” *Phys. Rev. A*, vol. 31, Feb 1985. [7.2](#)
- [106] A. O. Caldeira and A. J. Leggett, “Path integral approach to quantum Brownian motion,” *Physica A*, vol. 121, 1983. [7.2](#)
- [107] J. Allinger and U. Weiss, “Non-universality of dephasing in quantum transport,” *Zeitschrift für Physik B*, vol. 98, 1995. [7.2](#)
- [108] A. Venugopalan, D. Kumar, and R. Ghosh, “Environment-induced decoherence I. The Stern-Gerlach measurement,” *Physica A*, vol. 220, 1995. [7.2](#)
- [109] A. Venugopalan, D. Kumar, and R. Ghosh, “Environment-induced decoherence II. Effect of decoherence on Bell’s inequality for an EPR pair,” *Physica A*, vol. 220, 1995. [7.2](#)
- [110] G. Lindblad, “On the generators of quantum dynamical semigroups,” *Commun. Math. Phys.*, vol. 48, 1976. [7.2.1](#)
- [111] B. B. Mandelbrot and J. W. van Ness, “Fractional Brownian motions, fractional noises and applications,” *SIAM Review*, vol. 10, no. 4, 1968. [7.3](#)
- [112] B. Dupire, “Functional Itô calculus.” Portfolio Research Paper, Bloomberg, 2009. [7.3](#)
- [113] S. J. A. Malham and A. Wiese, “An introduction to SDE simulation,” *arXiv:1004.0646v1 [math.NA]*, April 2010. [A.1](#)
- [114] H. L. Anderson, ed., *A Physicist’s Desk Reference*. American Inst. of Phys., 1989. [A.3](#)
- [115] R. Gilmore, *Lie groups, physics, and geometry*. Cambridge Univ. Press, 2008. [B.2](#)

## VITA

Sam Kennerly received his BA degree in Mathematics from Northwestern University in Evanston, Illinois in 2001. He attended the MSc Theoretical Physics programme at the Universiteit van Amsterdam in The Netherlands for the 2004-5 academic year before leaving due to financial difficulties. In 2006 he enrolled in the PhD Physics program at Drexel University in Philadelphia, Pennsylvania.

### Publications

- S. Kennerly, “Illusory decoherence,” *Foundations of Physics*, vol. 42, no. 9, Sep 2012.
- J. Milutinovic, D. Pokrajac, and S. Kennerly, “GA optimization of point placement on a hypersphere,” *Intl. Conference on Mathematical Problems in Engineering, Aerospace and Science*, 2010.
- Z. Thraikill, S. Kennerly, and R. Ramos, “Modeling three and four coupled phase qubits,” *IEEE Transactions on Applied Superconductivity* 19, no. 3, Jun 2009.
- A. Tyler, R. Ramos, Z. Thraikill, and S. Kennerly, “Berry’s phase of a current-biased Josephson junction,” *Bulletin of the Amer. Phys. Soc.*, vol. 54, no. 1, Mar 2009.

While at Drexel, he also served as Station Manager and Chief Engineer of WKDU Philadelphia 91.7 FM, Drexel University’s free-format, noncommercial radio station.