Part I

Background

Chapter 1

Introduction to decoherence and noise in open quantum systems

Daniel A. Lidar and Todd A. Brun

1.1 Introduction

This chapter gives the physical and mathematical basis of decoherence in quantum systems, with particular emphasis on its importance in quantum information theory and quantum computation. We assume that the reader is already familiar with the basics of quantum mechanics, such as wave functions, the Schrödinger equation, and Hilbert spaces, which are covered in standard textbooks. Nevertheless, all basic concepts will be at least briefly defined and described. A working knowledge of graduate-level linear algebra and calculus is also assumed.

In quantum information processing, the term decoherence is often used loosely to describe *any* kind of noise that can affect a quantum system. This can include sources of noise that seem, qualitatively, to be quite different from each other. These noise sources include:

(i) Random driving forces from the environment (e.g., Brownian motion).
(ii) Interactions that produce entanglement between the system and the environment.
(iii) Statistical imprecision in the experimental controls on the system (e.g., timing errors, frequency fluctuations, etc.).

Remarkably, all of these sources of noise can be described by an almost identical mathematical framework, in which pure quantum states can evolve into mixed quantum states, and the quantum effects necessary for information processing – interference, entanglement, and reversible evolution – are disrupted.

The chapter begins by examining a highly schematic model of decoherence, involving systems of qubits, but uses this model to present the mathematical formalism used in describing more realistic systems as well. We will derive from this simple model the idea of the *completely positive trace-preserving map*, the most widely used description of quantum noise, and show

Quantum Error Correction, ed. Daniel A. Lidar and Todd A. Brun. Published by Cambridge University Press.

how it is equivalent to a stochastic error model. From this basic framework, we will also describe a continuous noise model: the quantum *master equation*.

After describing basic models of decoherence, we will discuss experimental methods of characterizing the ill effects. In the simplest models, decoherence can often be characterized by a few simple time scales: commonly T_1 (the *thermalization* time scale) and T_2 (the *dephasing* time scale). On a less simplified basis, one can try to determine an exact mathematical description of the quantum evolution, including the effects of noise, by *quantum process tomography*.

In the remainder of the chapter we will briefly cover several additional topics. While the simplest case of decoherence is the case where we are simply trying to preserve quantum information, more generally we wish to process that information as well, generally by applying unitary *quantum gates*. We will show how even a single dominant source of noise can manifest itself as different noise processes for different gates. We will compare our simple model to more realistic environments, and briefly discuss new issues that may arise, such as non-Markovian evolution and Zeno-like effects. We will also touch on Hamiltonian environment modes, which are more suitable descriptions for some approaches to error correction or suppression (especially dynamical decoupling).

Various methods have been proposed to combat the effects of decoherence in quantum information processing, and these are the subject of the rest of this book. They include straightforward attempts to cool and isolate the quantum system from its environment; techniques based on error-correcting codes; the use of decoherence-free subspaces and subsystems when the decoherence has certain symmetry properties; dynamical decoupling methods of error suppression, using rapid unitary pulses to suppress the effects of the system–environment interaction; and the use of topological properties to encode protected quantum information so that it is robust against local perturbations.

1.2 Brief introduction to quantum mechanics and quantum computing

There are many outstanding general introductions to quantum mechanics, and our purpose here is not to try to add to this list, but rather to provide a quick and relatively technical introduction to many of the concepts used later in the book. For a general introduction we recommend Chapter 2 in the excellent book by Nielsen and Chuang [NC00]. Our introduction to some extent overlaps with that book, but also complements it by covering a number of topics not mentioned there, which are particularly relevant to recent developments in quantum decoherence and noise mitigation. In particular, we cover the Dyson and Magnus expansions of time-dependent perturbation theory, general maps on quantum states, and weak measurements. Another excellent introduction is the book by Breuer and Petruccione [BP02], which covers the theory of open quantum systems in great detail.

1.2.1 States and spaces

Quantum states reside in, or act on, Hilbert spaces. A *Hilbert space* $\mathscr{H}$ is a real or complex inner product space (always complex in quantum mechanics) that is also a complete metric space with respect to the distance function induced by the inner product. A metric space M is a set where a notion of distance (called a metric) between elements of the set is defined. M is called complete if every Cauchy sequence in M converges in M.

1.2.1.1 Quantum states Quantum mechanics makes predictions about the evolution of quantum states and the outcomes of measurements when a system is in a particular quantum state. A general quantum state is a positive operator ρ called a *density matrix* whose trace is equal to 1, and whose domain and range are Hilbert spaces. We denote the set of all states by

$$\mathcal{S}(\mathcal{H}) \equiv \{\rho \in \mathcal{B}(\mathcal{H}) : \rho \geq 0, \mathrm{Tr}\rho = 1\}, \tag{1.1}$$

where $\mathcal{B}(\mathcal{H})$ is the set of all bounded linear operators on $\mathcal{H}$. "Bounded" means that the operator has a finite norm; we will define such norms later, in Section 1.2.6. A more general situation is when $\mathcal{B}$ maps between two different Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, in which case we write $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$. Of course, linear operators need not always be bounded (though in most of the applications in this book that will indeed be the case). We denote by $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ the set of all linear operators from Hilbert space $\mathcal{H}_1$ to Hilbert space $\mathcal{H}_2$, and by $\mathcal{T}(\mathcal{H}_1, \mathcal{H}_2)$ the set of all *superoperators*: mappings from $\mathcal{L}(\mathcal{H}_1)$ to $\mathcal{L}(\mathcal{H}_2)$ [the notation $\mathcal{L}(\mathcal{H})$ is shorthand for $\mathcal{L}(\mathcal{H}, \mathcal{H})$]. The name "superoperator" reflects the fact that a map $\Phi \in \mathcal{T}(\mathcal{H}_1, \mathcal{H}_2)$ transforms operators acting on $\mathcal{H}_1$ into operators acting on $\mathcal{H}_2$. Note that the two Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$ may have different dimensions.

An operator A is *Hermitian* if all its eigenvalues are real, or equivalently if $A = A^\dagger$, where $A^\dagger$ denotes the operation of taking the transpose of A as well the complex conjugate of every matrix element, whichever basis A is represented in. The notation $\rho \geq 0$ means that ρ is *positive semidefinite*, i.e., it is Hermitian and all of its eigenvalues are nonnegative. The sum of all of ρ's diagonal elements is denoted $\mathrm{Tr}\rho$.[1] Since $\mathrm{Tr}\rho = 1$, clearly at least one of the eigenvalues must be strictly positive.

The properties $\rho \geq 0$ and $\mathrm{Tr}\rho = 1$ are a consequence of the probabilistic interpretation of quantum mechanics, which states that a quantum state yields a valid probability distribution over possible measurement outcomes for any measurement. When ρ is written as a matrix in a basis in which it is diagonal, its diagonal elements are probabilities of mutually exclusive events. Therefore these diagonal elements must be positive and sum to one. When ρ has *rank* 1 (the rank is the number of linearly independent eigenvectors with nonzero eigenvalues) it is said to be *pure*. Otherwise it is *mixed*. A simple test for this, which one can easily verify, is a computation of $\mathrm{Tr}(\rho^2)$. When $\mathrm{Tr}(\rho^2) = 1$ the state is pure, and when $\mathrm{Tr}(\rho^2) < 1$ it is mixed.

When ρ is pure it can be written as an outer product of two vectors, which reads $\rho = |\psi\rangle\langle\psi|$ in standard Dirac bra-ket notation. When it is written in an explicit basis representation the normalized vector (ket) $|\psi\rangle \in \mathcal{H}$ is called the wave function in many textbooks. It is convenient to think operationally of a bra $\langle\psi|$ as the row vector that is the complex conjugate and transpose of the column vector ket $|\psi\rangle$, i.e., $\langle\psi| = |\psi\rangle^\dagger$.

Let us pick an orthonormal basis $\{|i\rangle\}_{i=0}^{d-1}$ for the Hilbert space $\mathcal{H}$, where $d \equiv \dim(\mathcal{H})$ is the dimension of the Hilbert space. We write the inner product in Dirac notation as $\langle\cdot|\cdot\rangle$. This is in general a complex number. Orthonormality of the basis is then expressed as $\langle i|j\rangle = \delta_{ij}$, where δ_{ij} is the Kronecker symbol ($\delta_{ij} = 1$ if $i = j$, and 0 otherwise). Since $\{|i\rangle\}$ is a basis we can write $|\psi\rangle = \sum_{i=0}^{d-1} \alpha_i |i\rangle$ for any ket $|\psi\rangle$, while the corresponding bra becomes

[1] More formally, in this book we deal almost exclusively with linear trace-class bounded operators that map between separable Hilbert spaces. A Hilbert space $\mathcal{H}$ is separable if and only if it admits a countable orthonormal basis. A bounded linear operator $A : \mathcal{H} \mapsto \mathcal{H}$, where $\mathcal{H}$ is separable, is said to be in the trace class if for some (and hence all) orthonormal bases $\{|k\rangle\}_k$ of $\mathcal{H}$ the sum of positive terms $\sum_k \langle k|\sqrt{A^\dagger A}|k\rangle$ is finite. In this case, the sum $\sum_k \langle k|A|k\rangle$ is absolutely convergent and is independent of the choice of the orthonormal basis. This value is called the trace of A, denoted by $\mathrm{Tr}(A)$.

$\langle\psi| = \sum_{i=0}^{d-1} \alpha_i^* \langle i|$. The numbers $\alpha_i = \langle i|\psi\rangle \in \mathbb{C}$ are called *amplitudes*. Since $|\psi\rangle$ is normalized we have $1 = \langle\psi|\psi\rangle = \sum_{i=0}^{d-1} |\alpha_i|^2$. This implies $\mathrm{Tr}(\rho) = 1$, since for a pure state $\mathrm{Tr}(\rho) = \langle\psi|\psi\rangle$. The positive numbers $|\alpha_i|^2 \leq 1$ play the role of a probability distribution, a fact we will explain in more detail when we discuss measurements, below.

The *outer product* of any two vectors $|\psi\rangle = \sum_{i=0}^{d-1} \alpha_i |i\rangle$ and $|\phi\rangle = \sum_{i=0}^{d-1} \beta_i |i\rangle$ can be written in the basis $\{|i\rangle\}$ as $|\psi\rangle\langle\phi| = \sum_{i,i'=0}^{d-1} \alpha_i \beta_{i'}^* |i\rangle\langle i'|$, where the operator $|i\rangle\langle i'| \in \mathscr{B}(\mathscr{H})$ acts as $(|i\rangle\langle i'|)|i''\rangle = \langle i'|i''\rangle|i\rangle$ on any basis element $|i''\rangle$.

1.2.1.2 Qubits The case $d = 2$ of a two-level system is special: it defines a quantum bit, or *qubit*, for which a pure state can always be written as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad |\alpha|^2 + |\beta|^2 = 1. \tag{1.2}$$

The basis states $|0\rangle$ and $|1\rangle$ are defined to be the eigenstates of the Pauli spin matrix σ_z with eigenvalues 1 and -1, respectively. Here are all three Pauli spin matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = X, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = Y, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = Z. \tag{1.3}$$

In addition, often the 2×2 identity matrix I is denoted by σ_0. A qubit is the simplest possible quantum mechanical system. There are many physical embodiments of such a system including the spin of a spin-1/2 particle (e.g., an electron), the polarization states of a photon, two hyperfine states of a trapped atom or ion, two neighboring levels of a Rydberg atom, or the presence or absence of a photon in a microcavity. All of these have been proposed in various schemes for quantum information and quantum computation, and used in actual experiments. The review [LJL+10] provides a comprehensive survey.

A global phase may be assigned arbitrarily to the state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ in Eq. (1.2), so that all physically distinct pure states of a single qubit form a two-parameter space. A useful parametrization is in terms of two angular variables θ and ϕ:

$$|\psi\rangle = \cos(\theta/2)e^{-i\phi/2}|0\rangle + \sin(\theta/2)e^{i\phi/2}|1\rangle \,, \tag{1.4}$$

where $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$. These two parameters define a point on the *Bloch sphere*. The north and south poles of the sphere represent the eigenstates of σ_z, and the eigenstates of σ_x and σ_y lie on the equator. Orthogonal states always lie opposite each other on the sphere.

If we allow states to be mixed, we represent a qubit by a density matrix ρ; the most general qubit density matrix can be written as

$$\rho = p|\psi\rangle\langle\psi| + (1-p)|\bar{\psi}\rangle\langle\bar{\psi}| \,, \tag{1.5}$$

where $|\psi\rangle$ and $|\bar{\psi}\rangle$ are two orthogonal pure states, $\langle\psi|\bar{\psi}\rangle = 0$. The mixed states of a qubit form a three parameter family:

$$\begin{aligned} \rho = & \left[\frac{1+r}{2}\cos^2(\theta/2) + \frac{1-r}{2}\sin^2(\theta/2)\right]|0\rangle\langle 0| \\ & + \left[\frac{1+r}{2}\sin^2(\theta/2) + \frac{1-r}{2}\cos^2(\theta/2)\right]|1\rangle\langle 1| \\ & + r\cos(\theta/2)\sin(\theta/2)\left[e^{i\phi}|0\rangle\langle 1| + e^{-i\phi}|1\rangle\langle 0|\right] \,, \end{aligned} \tag{1.6}$$

where θ and ϕ are the same angular parameters as before and $0 \leq r \leq 1$. The limit $r = 1$ is the set of pure states, parametrized as in Eq. (1.4), while $r = 0$ is the *completely mixed state* $\rho = I/2$, where $I = \sigma_0$ is the identity matrix. This state is the same as the state of a classical unbiased coin. Thus we can think of the Bloch sphere as having pure states on its surface and mixed states in its interior; and the distance r from the center is a measure of the state's purity. It is simply related to the parameter p in Eq. (1.5): $p = (1 + r)/2$.

For two qubits, the Hilbert space $\mathscr{H}_2 \otimes \mathscr{H}_2$ has a tensor-product basis:

$$|0\rangle_A \otimes |0\rangle_B \equiv |00\rangle, \tag{1.7a}$$
$$|0\rangle_A \otimes |1\rangle_B \equiv |01\rangle, \tag{1.7b}$$
$$|1\rangle_A \otimes |0\rangle_B \equiv |10\rangle, \tag{1.7c}$$
$$|1\rangle_A \otimes |1\rangle_B \equiv |11\rangle. \tag{1.7d}$$

We define tensor products formally in Section 1.2.4. Similarly, for N qubits we can define a basis $\{|j_{N-1}j_{N-2}\cdots j_0\rangle\}$, $j_k = 0, 1$. A useful labeling of these 2^N basis vectors is by the integers $0 \leq j < 2^N$ whose binary expressions are $j_{N-1}\cdots j_0$:

$$|j\rangle \equiv |j_{N-1}\cdots j_0\rangle, \quad j = \sum_{k=0}^{N-1} j_k 2^k \ . \tag{1.8}$$

1.2.1.3 Mixed states and the Bloch space Since a mixed state has rank at least 2, it can be written as

$$\rho = \sum_{i=1}^{d} r_i |\phi_i\rangle\langle\phi_i|, \tag{1.9}$$

where at least two of the r_is are nonzero and the set $\{|\phi_i\rangle\}$ is again an orthonormal basis for $\mathscr{H}$. This is called the *spectral decomposition* of ρ, since the r_i are its eigenvalues and the $|\phi_i\rangle$ are its eigenvectors. Since $\mathrm{Tr}(\rho) = 1$ we have $\sum_{i=1}^{d} r_i = 1$ and since $\rho \geq 0$ we have $r_i \geq 0$ for all i. If we pick an arbitrary orthonormal basis $\{|i\rangle\}$ for $\mathscr{H}$ then ρ need not be diagonal and can be written in the form $\rho = \sum_{i=1}^{d} a_{ij}|i\rangle\langle j|$. The matrix $A \equiv \{a_{ij}\}$ then satisfies $\mathrm{Tr}(A) = 1$ and $A \geq 0$. Every positive matrix can be diagonalized with a unitary transformation. When this is done we recover the spectral decomposition (1.9).

It is possible to give a geometric characterization of density matrices by using the *Bloch vector representation* for an arbitrary d-dimensional Hilbert space $\mathscr{H}$ [KK05]. This works as follows: Let $F_0 \equiv I$ and let $\{F_\mu : \mathrm{Tr}(F_\mu) = 0\}_{\mu=1}^{d^2-1}$ be a basis for the set of traceless Hermitian matrices in $\mathscr{B}(\mathscr{H})$. Assume further that $\mathrm{Tr}(F_\mu F_\nu) = d\delta_{\mu\upsilon}$ $\forall \mu, \nu$, i.e., the Fs are mutually orthogonal with respect to the *Hilbert–Schmidt inner product* $\langle A|B\rangle \equiv \mathrm{Tr}(A^\dagger B)$, which is the natural inner product on $\mathscr{B}(\mathscr{H})$, when this space is a Hilbert space. Hence any state ρ can be expanded as

$$\rho = \frac{1}{d}\left(I + \sum_{\mu=1}^{d^2-1} b_\mu F_\mu\right); \quad b_\mu = \mathrm{Tr}[\rho F_\mu] \equiv \langle F_\mu\rangle_\rho, \tag{1.10}$$

The vector $\mathbf{b} = (b_1, ..., b_{N^2-1}) \in \mathbb{R}^{N^2-1}$ of expectation values (more on this notion in the measurements, Section 1.2.5) is called the *Bloch vector*, and knowing its components is equivalent to complete knowledge of the corresponding state ρ, via the mapping $\mathbf{b} \mapsto \rho = \frac{1}{d}(I + \sum_{\mu=1}^{d^2-1} b_\mu F_\mu)$.

Let $\mathbf{n}$ denote a unit vector, i.e., $\mathbf{n} \in \mathbb{R}^{d^2-1}$ and $\sum_{i=1}^{d^2-1} n_i^2 = 1$, and define $F_{\mathbf{n}} \equiv \sum_{\mu=1}^{d^2-1} n_\mu F_\mu$. Let the minimum eigenvalue of each $F_{\mathbf{n}}$ be denoted $m(F_{\mathbf{n}})$. The "Bloch space" $\boldsymbol{B}(\mathbb{R}^{d^2-1})$ is the set of all Bloch vectors and is a closed convex set, since the set of states $\mathscr{S}(\mathscr{H})$ is closed and convex, and the map $\mathbf{b} \mapsto \rho$ is linear homeomorphic. The Bloch space is characterized in the "spherical coordinates" determined by $\{F_{\mathbf{n}}\}$ as

$$\boldsymbol{B}(\mathbb{R}^{d^2-1}) = \left\{ \mathbf{b} = r\mathbf{n} \in \mathbb{R}^{d^2-1} : r \leq \frac{1}{|m(F_{\mathbf{n}})|} \right\}. \tag{1.11}$$

This result is useful for visualization of quantum states. For example, for two qubits the Bloch space is given by Eq. (1.11) with $d = 4$, which corresponds to a certain 15-dimensional convex set. The Bloch space of a qubit is defined with the $\{F_\mu\}$ being the Pauli matrices; it is a simple sphere, since it so happens that for a qubit the minimum eigenvalues $m(F_{\mathbf{n}})$ are 1 for all $F_{\mathbf{n}}$.

1.2.2 Quantum gates and the dynamics of isolated (closed) systems

1.2.2.1 Schrödinger equation The evolution of a quantum state is governed by the Schrödinger equation:

$$\frac{\partial \rho}{\partial t} = -\frac{i}{\hbar}[H, \rho]. \tag{1.12}$$

The Hermitian operator $H \in \mathscr{B}(\mathscr{H})$ is called the Hamiltonian and is the "generator" of the evolution. It is the sum of the (quantized) kinetic and potential energies of the system. The symbol $\hbar \approx 10^{-34}$ joule second is the Planck constant and has units of energy multiplied by time (or, equivalently, units of action). From now on we will use units where $\hbar = 1$, which means that H has units of inverse time (frequency). The symbol $[\cdot, \cdot]$ is the commutator: $[H, \rho] \equiv H\rho - \rho H$. Since H is Hermitian the solution to the Schrödinger equation is

$$\rho(t) = U(t)\rho(0)U^\dagger(t), \tag{1.13}$$

where U is a *unitary* operator ($U^\dagger = U^{-1}$) sometimes called the propagator, and is related to H via

$$\frac{\partial U}{\partial t} = -iHU. \tag{1.14}$$

When H is time independent we have the simple solution $U(t) = \exp(-iHt)$, assuming the boundary condition $U(0) = I$, where I is the identity operator.

Note that given $U(t)$ for all times t we can compute the Hamiltonian as the logarithm of $U(t)$. However, rather than taking the actual log of $U(t)$, which would require proper handling of branch cuts in the complex plane, we can obtain $H(t)$ by computing the "logarithmic derivative":

$$H(t) = -i\frac{\partial U(t)}{\partial t}U^{-1}(t). \tag{1.15}$$

1.2.2.2 Dyson series When H is time dependent it need not commute with itself at different times and Eq. (1.14) has the formal solution

$$U(t) = T_+ \exp\left[-i\int_0^t H(t')dt'\right], \tag{1.16}$$

where T_+ is the *time-ordering operator*. This means that the solution can be expressed as an infinite sum, called the *Dyson series*:

$$U(t) = I + \sum_{n=1}^{\infty} S_n(t), \tag{1.17}$$

where each term is a time-ordered multiple integral

$$S_n(t) \equiv (-i)^n \int_0^t dt_1 H(t_1) \int_0^{t_1} dt_2 H(t_2) \cdots \int_0^{t_{n-1}} dt_n H(t_n), \tag{1.18}$$

where "time-ordered" means with early to late times being composed from right to left.

When H is time independent each of the multiple integrals becomes $S_n(t) = (-iHt)^n/n!$, so that $U(t) = \exp(-iHt)$, as it should.

The Dyson series (1.17) has the advantage of providing a simple recipe for generating terms to arbitrarily high order. As we will see in Section 1.2.6.1, it also converges for all t provided H is a bounded operator for all t, but it has the disadvantage that the partial sum $U_k(t) \equiv I + \sum_{n=1}^{k} S_n(t)$ is not unitary. This is sometimes a problem in quantum computation applications, where we are often interested in unitary transformations.

1.2.2.3 Magnus expansion An alternative to the Dyson series is the *Magnus expansion*, which has the advantage that it is unitary to all orders. This makes it particularly useful for applications in quantum computation, such as in dynamical decoupling, as we will see in Chapter 4. Like the Dyson expansion, the Magnus expansion at time t is an operator series

$$\Omega(t) \equiv \sum_{n=1}^{\infty} \Omega_n(t) \tag{1.19}$$

such that $\Omega_n(t)$ is nth order in the Hamiltonian $H(t)$, but now the series appears in the exponential:

$$U(t) = \exp\left[\Omega(t)\right]. \tag{1.20}$$

Thus, for a fixed time T, the time evolution generated by the time-dependent Hamiltonian $H(t)$ is equivalent to the time evolution generated by the time-independent effective Hamiltonian $H_{\text{eff}} \equiv \frac{i}{T}\Omega(T)$.

The first few terms in the Magnus expansion are

$$\Omega_1(t) = -i \int_0^t dt_1\, H(t_1), \tag{1.21a}$$

$$\Omega_2(t) = -\frac{1}{2} \int_0^t dt_1 \int_0^{t_1} dt_2 \left[H(t_1), H(t_2)\right], \tag{1.21b}$$

$$\Omega_3(t) = \frac{i}{6} \int_0^t dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3 ([H(t_1), [H(t_2), H(t_3)]] + [H(t_3), [H(t_2), H(t_1)]]). \tag{1.21c}$$

Higher-order terms can be computed using a recursive formula, but unfortunately, in contrast to the Dyson series, they do not have a simple structure. In general, $\Omega_n(t)$ is the time integral of a sum of $(n-1)$-nested commutators with coefficients related to the Bernoulli numbers, each with

n factors of $H(t)$. The Magnus expansion is thus an infinite series in Ht; a sufficient condition for convergence is

$$\int_0^t dt' \ \|H(t')\| < \pi, \tag{1.22}$$

but the Magnus expansion need not otherwise converge, which is a disadvantage it has compared to the Dyson expansion. The review [BCO+09] provides a comprehensive discussion of the Magnus and Dyson expansions.

1.2.2.4 Quantum gates You can verify that when $\rho = |\psi\rangle\langle\psi|$, i.e., a pure state, the Schrödinger equation (1.12) becomes the familiar

$$\frac{\partial|\psi\rangle}{\partial t} = -iH|\psi\rangle, \tag{1.23}$$

and has the solution $|\psi(t)\rangle = U(t)|\psi(0)\rangle$. The propagator $U(t)$ is the same unitary operator as the one we just discussed.

It is often convenient to treat time evolutions at the level of unitary transformations rather than explicitly solving the Schrödinger equation. In such cases time can be treated as a discrete variable:

$$|\psi_n\rangle = U_n U_{n-1} \cdots U_1 |\psi_0\rangle. \tag{1.24}$$

If the unitary operator U_n is *weak*, that is, close to the identity, one can always find a Hamiltonian operator H_n such that

$$U_n = e^{-iH_n \delta t} \approx I - iH_n \delta t, \tag{1.25}$$

where δt is an appropriately short time interval. Thus, one can easily recover the Schrödinger equation from a description in terms of unitary operators:

$$\delta|\psi_n\rangle = |\psi_n\rangle - |\psi_{n-1}\rangle = (U_n - I)|\psi_{n-1}\rangle \approx -iH_n|\psi_{n-1}\rangle\delta t \ . \tag{1.26}$$

Linear combinations of the Pauli operators σ_x, σ_y, and σ_z, together with the identity I, are sufficient to produce any operator on a single qubit. To specify any unitary transformation it suffices to give its effect on a complete set of basis vectors. We will consider only a fairly limited set of two-qubit transformations, and no transformations involving more than two qubits, but the simple formalism we derive readily generalizes to higher-dimensional systems.

Let us examine a couple of examples of two-qubit transformations. The controlled-NOT gate (or CNOT) is widely used in quantum computation; applied to the tensor-product basis vectors [Eqs. (1.7a)–(1.7d)] it gives

$$U_{\rm CNOT}|00\rangle = |00\rangle, \tag{1.27a}$$
$$U_{\rm CNOT}|01\rangle = |01\rangle, \tag{1.27b}$$
$$U_{\rm CNOT}|10\rangle = |11\rangle, \tag{1.27c}$$
$$U_{\rm CNOT}|11\rangle = |10\rangle \ . \tag{1.27d}$$

If the first qubit is in state $|0\rangle$, this gate leaves the second qubit unchanged; if the first qubit is in state $|1\rangle$, the second qubit is flipped $|0\rangle \leftrightarrow |1\rangle$. Hence the name: whether a NOT gate is performed on the second qubit is *controlled* by the first qubit. In terms of single-qubit operators, $U_{\rm CNOT} = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes \sigma_x$.