Quantum Notes

Henry Yu

Contents

Chapter 1

Single Systems

1.1 Introduction to Quantum Information

Definition 1.1.1 (Simplified Description)

Quantum states are represented by vectors with operations represented by matrices. This is sufficient for understanding most quantum algorithms.

Definition 1.1.2 (General Description)

Quantum states are represented by density matrices with operations represented by superoperators. This definition allows for a more general class of measurements and operations. The simplified description and classical information (including probabilitstic states) are special cases of this definition.

1.2 Classical Information

Consider a physical system that stores information: let us call it X. Assume X can be in one of a finite number of classical states at each moment. Denote this classical state set by Σ.

Example 1.2.1

- If X is a bit, then its classical state set is $\Sigma = \{0, 1\}$.
- If X is a six-sided die, then $\Sigma = \{1, 2, 3, 4, 5, 6\}$.
- If X is a switch on a fan, then $\Sigma = \{\text{high, medium, low, off}\}$.

There has to be at least one classical state that the system can be in (which means no empty set) and at least two if it is going to be useful for storing information.

Often in information processing, our knowledge of the state of X is a bit uncertain; therefore, we can assign probabilities to each of its possible classical states. For example, if X is a bit, then perhaps it is in the state 0 with probability $\frac{3}{4}$ and in the state 1 with probability $\frac{1}{4}$. This is a **probabilistic state** of X.

Example 1.2.2 (Probabilistic states of X)

$$\Pr(X = 0) = \frac{3}{4} \quad \text{and} \quad \Pr(X = 1) = \frac{1}{4}$$

Another succinct way to represent this is by a **column vector** (probability vector):

Example 1.2.3 (Column vector of X)

$$\begin{pmatrix} \frac{3}{4} \\ \frac{1}{4} \end{pmatrix} \quad \begin{matrix} \text{entry corresponding to } 0 \\ \text{entry corresponding to } 1 \end{matrix}$$

Remark All entries of a classical probabilistic vector are nonnegative real numbers and sum to 1.

1.2.1 Dirac notation (pt.1)

Let Σ be any classical state set, and assume the elements of Σ have been placed in correspondence with the integers $1, \cdots, |\Sigma|$.

We denote $|a\rangle$ ("ket a") as the column vector having a 1 in the entry corresponding to $a \in \Sigma$, with 0 for all other entries.

Example 1.2.4

If $\Sigma = \{0, 1\}$, then $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

If $\Sigma = \{\clubsuit, \diamond, \heartsuit, \spadesuit\}$ and we assume this ordering, this yields:

- $|\clubsuit\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$
- $|\diamond\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$
- $|\heartsuit\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$
- $|\spadesuit\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$

Note:

Vectors of this form are called standard basis vectors. Every vector can be expressed uniquely as a linear combination of standard basis vectors. For example, $\begin{pmatrix} \frac{3}{4} \\ \frac{1}{4} \end{pmatrix} = \frac{3}{4}|0\rangle + \frac{1}{4}|1\rangle$.

If you happen to measure the state of X being $a \in \Sigma$, then $\Pr(X = a) = 1$. Measuring X selects a transition, chosen at random. For example, using $\frac{3}{4}|0\rangle + \frac{1}{4}|1\rangle$, X can be measured as $|0\rangle$ with probability $\frac{3}{4}$ or $|1\rangle$ with probability $\frac{1}{4}$.

1.2.2 Deterministic operations

There's no element of chance when a deterministic operation is performed, so there's no randomness or uncertainty involved.

Every function $f : \Sigma \to \Sigma$ describes a deterministic operation that transforms the classical state a into $f(a)$, for each $a \in \Sigma$.

Definition 1.2.1 (Matrix Representation)

Given any function $f : \Sigma \to \Sigma$, there is a unique matrix M satisfying $M|a\rangle = |f(a)\rangle$ for every $a \in \Sigma$. This matrix has exactly one 1 in each column, and 0 for all other entries:

$$M(b,a) = \begin{cases} 1 & b = f(a) \\ 0 & b \neq f(a) \end{cases}$$

Example 1.2.5

For $\Sigma = \{0,1\}$, there are four functions of the form $f : \Sigma \to \Sigma$:

a	$f_1(a)$
0	0
1	0

zero function
$f_1(0) = 0$
$f_1(1) = 0$

a	$f_2(a)$
0	0
1	1

identity function
$f_2(0) = 0$
$f_2(1) = 1$

a	$f_3(a)$
0	1
1	0

NOT function
$f_3(0) = 1$
$f_3(1) = 0$

a	$f_4(a)$
0	1
1	1

one function
$f_4(0) = 1$
$f_4(1) = 1$

Here are the matrices corresponding to these functions:

$$M_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad M_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad M_4 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$$

Note:

Recall that the 0 state is represented by the column vector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. and the 1 state is represented by the column vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

1.2.3 Dirac notation (pt.2)

Let Σ be any classical state set, and assume the elements of Σ have been placed in correspondence with the integers $1, \cdots, |\Sigma|$.

We denote by $\langle a|$ ("bra a") the row vector having a 1 in the entry corresponding to $a \in \Sigma$, with 0 for all other entries.

If $\Sigma = \{0,1\}$, then $\langle 0| = \begin{pmatrix} 1 & 0 \end{pmatrix}$ and $\langle 1| = \begin{pmatrix} 0 & 1 \end{pmatrix}$.

Multiplying a row vector to a column vector yields a scalar:

$$\begin{pmatrix} * & * & \cdots & * \end{pmatrix} \begin{pmatrix} * \\ * \\ \vdots \\ * \end{pmatrix} = (*)$$

$$\langle a|b\rangle = \begin{cases} 1 & a = b \\ 0 & a \neq b \end{cases}$$

Multiplying a column vector to a row vector yields a matrix:

$$\begin{pmatrix} * \\ * \\ \vdots \\ * \end{pmatrix} \begin{pmatrix} * & * & \cdots & * \end{pmatrix} = \begin{pmatrix} * & * & \cdots & * \\ * & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \end{pmatrix}$$

In general, $|a\rangle\langle b|$ has a 1 in the (a,b)-entry and 0 for all other entries.

The matrix may now be expressed as $M = \sum_{b \in \sum} |f(b)\rangle\langle b|$.

1.2.4 Probabilistic operations

Probabilistic operations are classical operations that may introduce randomness or uncertainty. They are described by stochastic matrices:

- All entries are nonnegative real numbers.
- All columns sum to 1.

Example 1.2.6 (Probabilistic operation on a bit)

- If the classical state is 0, then do nothing.
- If the classical state is 1, then flip the bit with probability $\frac{1}{2}$.

Here it is visualized as a matrix:

$$\begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

1.2.5 Composing operations

Example 1.2.7

Suppose X is a system and $M_1, \cdots, M_n$ are stochastic matrices representing probabilistic operations on X. Applying the first probabilistic operation to the probability vecotr v, then applying the second probabilistic operation to the result yields this vector:

$$M_2(M_1 v) = (M_2 M_1)v$$

The probabilistic operation obtained by composing the first and second probabilistic operations is represented by the matrix product M_2M_1. Composing the probabilistic operations represented by the matrices $M_1, \cdots, M_n$ is represented by the matrix product:

$$M_n \cdots M_2 M_1$$

Remark The order is important: matrix multiplication is **not** commutative.

1.3 Quantum Information

A quantum state of a system is represented by a column vector whose indices are places in correspondence with the classical states of that system:

- The entries are complex numbers.
- The sum of the absolute values squared of the entries is 1.

The Euclidean norm (or ℓ^2-norm) for vectors with complex entries is defined as:

$$v = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \Rightarrow ||v|| = \sqrt{\sum_{k=1}^{n} |\alpha_k|^2}$$

Quantum state vectors are therefore unit vectors with respect to this norm.

Example 1.3.1 (Qubit states)

- Standard basis states: $|0\rangle$ and $|1\rangle$
- Plus/minus states:

$$|+\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle \text{ and } |-\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$$

- $\frac{1+2i}{3}|0\rangle - \frac{2}{3}|1\rangle$ (random state)

1.3.1 Dirac notation (pt.3)

The Dirac notation can be used for arbitrary vectors: any name can be used in place of a classical state.

Example 1.3.2

- Ket = $|\psi\rangle$ = column vectors
- Bra = $\langle\psi|$ = row vectors

The notation $|\psi\rangle$ (psi) is commonly used to refer to an arbitrary vector:

Example 1.3.3

- $|\psi\rangle = \frac{1+2i}{3}|0\rangle - \frac{2}{3}|1\rangle = \begin{pmatrix} \frac{1+2i}{3} \\ -\frac{2}{3} \end{pmatrix}$
- $\langle\psi| = \frac{1-2i}{3}\langle 0| - \frac{2}{3}\langle 1| = \begin{pmatrix} \frac{1-2i}{3} & -\frac{2}{3} \end{pmatrix}$

Note:

For any column vector $|\psi\rangle$, the corresponding row vector is $\langle\psi| = |\psi\rangle^\dagger$ (psi dagger), where $\dagger$ denotes the conjugate transpose. The notation $|\psi\rangle^\dagger$ means to take the transpose of $|\psi\rangle$ and then take the complex conjugate of each entry, or vice versa; the order does not matter.

1.3.2 Measuring quantum states

Things to note about measuring quantum states:

- The possible outcomes are the classical states.
- The probability for each classical state to be the outcome is the absolute value squared of the corresponding quantum state vector entry.

Measuring the quantum state $|0\rangle$ gives the outcome 0 with certainty, and measuring the quantum state $|1\rangle$ gives the outcome 1 with certainty.

Example 1.3.4

Measuring the quantum state $|+\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ yields an outcome as follows:

$$\Pr(\text{outcome is } 0) = |\frac{1}{\sqrt{2}}|^2 = \frac{1}{2} \qquad \Pr(\text{outcome is } 1) = |\frac{1}{\sqrt{2}}|^2 = \frac{1}{2}$$

Measuring a system changes its quantum state: if we obtain the classical state α, then the quantum state of the system is $|\alpha\rangle$.

Example 1.3.5

The system $\frac{1+2i}{3}|0\rangle - \frac{2}{3}|1\rangle$ has a probability of $\frac{5}{9}$ to "collapse" to the quantum state $|0\rangle$ and a probability of $\frac{4}{9}$ to "collapse" to the quantum state $|1\rangle$.

1.3.3 Unitary operations

Operations on quantum state vectors are represented by unitary matrices.

A square matrix $\mathcal{U}$ having complex entries is unitary if it satisfies the equalities

$$\mathcal{U}^\dagger\mathcal{U} = (\mathbb{1} \text{ or } I_n) = \mathcal{U}\mathcal{U}^\dagger$$

where $\mathcal{U}^\dagger$ is the conjugate transpose of $\mathcal{U}$ and $\mathbb{1}$ is the identity matrix.

The condition that an $n \times n$ matrix $\mathcal{U}$ is unitary is equivalent to $||\mathcal{U}v|| = ||v||$ for every n-dimensional column vector v with complex number entries.

If v is the quantum state vector, then $\mathcal{U}v$ is also a quantum state vector.

1.3.4 Qubit unitary operations

Example 1.3.6 (Pauli operations)

Pauli operators are ones represented by the Pauli matrices:

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

They also happen to be equal to their conjugate transposes, so they are also Hermitian matrices.
Common alternative notations: $\mathcal{X} = \sigma_x$, $\mathcal{Y} = \sigma_y$, $\mathcal{Z} = \sigma_z$

The operation σ_x is also called a bit flip (or a NOT operation) and the σ_z operation is called a phase flip:

$$\begin{aligned} \sigma_x |0\rangle &= |1\rangle & \qquad \sigma_z |0\rangle &= |0\rangle \\ \sigma_x |1\rangle &= |0\rangle & \qquad \sigma_z |1\rangle &= -|1\rangle \end{aligned}$$

Example 1.3.7 (Hadamard operation)

The Hadamard operation is represented by this matrix:

$$\mathcal{H} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Checking that $\mathcal{H}$ is indeed unitary:

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}^\dagger \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Seeing that the conjugate transpose is equal to itself, it also happens to be a Hermitian matrix.

Example 1.3.8 (Phase operations)

A phase operation is one described by this matrix:

$$\mathcal{P}_\theta = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}$$

with θ being any real number.

The operations

$$\mathcal{S} = \mathcal{P}_{\frac{\pi}{2}} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \text{ and } \mathcal{T} = \mathcal{P}_{\frac{\pi}{4}} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1+i}{\sqrt{2}} \end{pmatrix}$$

are important examples.

1.3.5 Composing unitary operations

Composition of unitary operations are represented by matrix multiplication (similar to the probabilistic setting), i.e., you can compute the action of a sequence of unitary operations just by multiplying the matrices together, which will end up being a single unitary matrix.

Chapter 2

Multiple Systems

2.1 Classical Information

2.1.1 Classical states

Suppose that we have two systems:

- X is a system having classical state set Σ.
- Y is a system having classical state set Γ.

Imagine that X and Y are placed side-by-side, with X on the left and Y on the right, and viewed together as if they form a single system.

We denote this new compound system by (X, Y), or XY.

Question 2.1.1
What are the classical states of (X, Y)?

Solution.
The classical state set of (X, Y) is the **Cartesian product**:

$$\Sigma \times \Gamma = \{(x, y) \mid x \in \Sigma, y \in \Gamma\}$$

Remark It's not important that X is on the left and Y is on the right; we could have just as well placed Y on the left and X on the right. The important thing is that the two systems are distinguishable.

Example 2.1.1 (Card suits)
If $\Sigma = \{0, 1\}$ and $\Gamma = \{\clubsuit, \diamond, \heartsuit, \spadesuit\}$, then:

$$\Sigma \times \Gamma = \{(0, \clubsuit), (0, \diamond), (0, \heartsuit), (0, \spadesuit), (1, \clubsuit), (1, \diamond), (1, \heartsuit), (1, \spadesuit)\}$$

This description generalizes to more than two systems in a natural way.

Suppose $\mathsf{X}_1, \cdots, \mathsf{X}_n$ are systems having classical state sets $\Sigma_1, \cdots, \Sigma_n$, respectively.

The classical state set of the n-tuple $(\mathsf{X}_1, \cdots, \mathsf{X}_n)$, viewed as a single compound system, is the Cartesian product:

$$\Sigma_1 \times \cdots \times \Sigma_n = \{(x_1, \cdots, x_n) \mid x_1 \in \Sigma_1, \cdots, x_n \in \Sigma_n\}$$

Example 2.1.2

If $\Sigma_1 = \Sigma_2 = \Sigma_3 = \{0, 1\}$, then the classical state set of $(\mathsf{X}_1, \mathsf{X}_2, \mathsf{X}_3)$ is:

$$\Sigma_1 \times \Sigma_2 \times \Sigma_3 = \{(0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1)\}$$

An n-tuple $(\mathsf{x}_1, \cdots, \mathsf{x}_n)$ may also be written as a string $\mathsf{x}_1 \cdots \mathsf{x}_n$.

Example 2.1.3 (Binary alphabet)

Suppose $\mathsf{X}_1, \cdots, \mathsf{X}_n$ are bits, so thei classical state sets are all the same:

$$\Sigma_1 = \Sigma_2 = \cdots = \Sigma_{10} = \{0, 1\}$$

The classical state set of $(\mathsf{X}_1, \cdots, \mathsf{X}_n)$ is the Cartesian product:

$$\Sigma_1 \times \Sigma_2 \times \cdots \times \Sigma_{10} = \{0, 1\}^{10}$$

You can also think about this as a 10-bit register in a classical computer. Written as strings, these classical states look like this:

0000000000
0000000001
0000000010
0000000011
⋮
1111111111

Note:

Cartesian products of classical state sets are ordered lexicographically (i.e., in dictionary order):

- We assume the individual classical state sets are already ordered.
- Significance decreases from left to right.

Example 2.1.4

The Cartesian product $\{1, 2, 3\} \times \{0, 1\}$ is ordered like this:

$$(1,0), (1,1), (2,0), (2,1), (3,0), (3,1)$$

When n-tuples are written as strings and ordered in this way, we observe familiar patterns, such as $\{0, 1\} \times \{0, 1\}$ being ordered as $00, 01, 10, 11$.

2.1.2 Probabilistic states

Probabilistic states of compound systems associates probabilities with the Cartesian product of the classical state sets of individual systems.

Example 2.1.5

This is a probabilistic state pair of bits (X, Y):

$$Pr((\mathsf{X}, \mathsf{Y}) = (0,0)) = \frac{1}{2}$$

$$Pr((\mathsf{X}, \mathsf{Y}) = (0,1)) = 0$$

$$Pr((\mathsf{X},\mathsf{Y}) = (1,0)) = 0$$

$$Pr((\mathsf{X},\mathsf{Y}) = (1,1)) = \frac{1}{2}$$

An alternate notation using vector notation:

$$\begin{pmatrix} \frac{1}{2} \\ 0 \\ 0 \\ \frac{1}{2} \end{pmatrix} \begin{array}{l} \leftarrow \text{ probability associated with state 00} \\ \leftarrow \text{ probability associated with state 01} \\ \leftarrow \text{ probability associated with state 10} \\ \leftarrow \text{ probability associated with state 11} \end{array}$$

For a given probabilistic state of (X,Y), we say that X and Y are **statistically independent** if

$$Pr((\mathsf{X},\mathsf{Y}) = (\mathsf{x},\mathsf{y})) = Pr(\mathsf{X} = \mathsf{x})Pr(\mathsf{Y} = \mathsf{y})$$

for all $\mathsf{x} \in \Sigma$ and $\mathsf{y} \in \Gamma$.

Suppose that a probabilistic state of (X,Y) is expressed as a vector:

$$|\pi\rangle = \sum_{(\mathsf{x},\mathsf{y})\in\Sigma\times\Gamma} p_{ab}\,|\mathsf{x}\mathsf{y}\rangle$$

The systems X and Y are independent if there exist probability vectors

$$|\phi\rangle = \sum_{\mathsf{x}\in\Sigma} q_x\,|\mathsf{x}\rangle \quad \text{and} \quad |\psi\rangle = \sum_{\mathsf{y}\in\Gamma} r_y\,|\mathsf{y}\rangle$$

such that $p_{\mathsf{x}\mathsf{y}} = q_x r_y$ for all $\mathsf{x} \in \Sigma$ and $\mathsf{y} \in \Gamma$.

Example 2.1.6 (Independent bits)
The probabilistic state of a pair of bits (X,Y) represented by the vector

$$|\pi\rangle = \frac{1}{6}|00\rangle + \frac{1}{12}|01\rangle + \frac{1}{2}|10\rangle + \frac{1}{4}|11\rangle$$

is one in which X and Y are independent. The required condition is true for these probability vectors:

$$|\phi\rangle = \frac{1}{4}|0\rangle + \frac{3}{4}|1\rangle \quad \text{and} \quad |\psi\rangle = \frac{2}{3}|0\rangle + \frac{1}{3}|1\rangle$$

Example 2.1.7 (Dependent bits)
For the probabilistic state

$$\frac{1}{2}|00\rangle + \frac{1}{2}|11\rangle$$

of two bits (X,Y), we have that X and Y are not independent.

If they were, we would have numbers q_0, q_1, r_0, r_1 such that

$$q_0 r_0 = \frac{1}{2}$$

$$q_0 r_1 = 0$$

$$q_1 r_0 = 0$$

$$q_1 r_1 = \frac{1}{2}$$

But if $q_0 r_1 = 0$, then either $q_0 = 0$ or $r_1 = 0$ (or both), contradicting either the first or last equality since any number multiplied by 0 is 0.

2.1.3 Tensor products of vectors

The tensor product of two vectors

$$|\phi\rangle = \sum_{x \in \Sigma} \alpha_x |x\rangle \quad \text{and} \quad |\psi\rangle = \sum_{y \in \Gamma} \beta_y |y\rangle$$

is the vector

$$|\phi\rangle \otimes |\psi\rangle = \sum_{(x,y) \in \Sigma \times \Gamma} \alpha_x \beta_y |xy\rangle$$

Equivalently, the vector $|\phi\rangle \otimes |\psi\rangle$ is defined by this condition:

$$\langle xy \mid \pi \rangle = \langle x \mid \phi \rangle \langle y \mid \psi \rangle \qquad (\text{for all } x \in \Sigma \text{ and } y \in \Gamma)$$

Remark In essense, this is the same operation for probabilistic states of a pair of independent bits; we are just giving a name to it now.

Following our convention for ordering the elements of Cartesian product sets, we obtain this specification for the tensor product of two column vectors:

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} \otimes \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} = \begin{pmatrix} \alpha_1\beta_1 \\ \vdots \\ \alpha_1\beta_k \\ \alpha_2\beta_1 \\ \vdots \\ \alpha_2\beta_k \\ \vdots \\ \alpha_m\beta_1 \\ \vdots \\ \alpha_m\beta_k \end{pmatrix}$$

Example 2.1.8 (Tensor product)

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \otimes \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix} = \begin{matrix} \alpha_1 \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix} \\ \alpha_2 \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix} \\ \alpha_3 \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix} \end{matrix} = \begin{matrix} \begin{pmatrix} \alpha_1\beta_1 \\ \alpha_1\beta_2 \\ \alpha_1\beta_3 \\ \alpha_1\beta_4 \end{pmatrix} \\ \begin{pmatrix} \alpha_2\beta_1 \\ \alpha_2\beta_2 \\ \alpha_2\beta_3 \\ \alpha_2\beta_4 \end{pmatrix} \\ \begin{pmatrix} \alpha_3\beta_1 \\ \alpha_3\beta_2 \\ \alpha_3\beta_3 \\ \alpha_3\beta_4 \end{pmatrix} \end{matrix} = \begin{pmatrix} \alpha_1\beta_1 \\ \alpha_1\beta_2 \\ \alpha_1\beta_3 \\ \alpha_1\beta_4 \\ \alpha_2\beta_1 \\ \alpha_2\beta_2 \\ \alpha_2\beta_3 \\ \alpha_2\beta_4 \\ \alpha_3\beta_1 \\ \alpha_3\beta_2 \\ \alpha_3\beta_3 \\ \alpha_3\beta_4 \end{pmatrix}$$

Observe the following expression for tensor products of standard basis vectors:

$$|x\rangle \otimes |y\rangle = |x\rangle |y\rangle = |xy\rangle$$

Alternatively, writing (x, y) as an ordered pair rather than a string, we could write

$$|x\rangle \otimes |y\rangle = |(x, y)\rangle$$

but it is more common to write

$$|x\rangle \otimes |y\rangle = |x, y\rangle$$

Note:

The tensor product of two vectors is **bilinear**.

1. Linearity in the first argument:

$$(|\phi_1\rangle + |\phi_2\rangle) \otimes |\psi\rangle = |\phi_1\rangle \otimes |\psi\rangle + |\phi_2\rangle \otimes |\psi\rangle$$

$$(\alpha |\phi\rangle) \otimes |\psi\rangle = \alpha(|\phi\rangle \otimes |\psi\rangle)$$

2. Linearity in the second argument:

$$|\phi\rangle \otimes (|\psi_1\rangle + |\psi_2\rangle) = |\phi\rangle \otimes |\psi_1\rangle + |\phi\rangle \otimes |\psi_2\rangle$$

$$|\phi\rangle \otimes (\alpha |\psi\rangle) = \alpha(|\phi\rangle \otimes |\psi\rangle)$$

Notice that scalars "float freely" within tensor products:

$$\alpha |\phi\rangle \otimes |\psi\rangle = \alpha(|\phi\rangle \otimes |\psi\rangle) = |\psi\rangle \otimes \alpha |\phi\rangle$$

There's no difference multiplying the first argument by the scalar α or multiplying the second argument by the scalar α.

Tensor products generalize to three or more systems.

If $|\phi_1\rangle, \cdots, |\phi_n\rangle$ are vectors, then the tensor product

$$|\psi\rangle = |\phi_1\rangle \otimes \cdots \otimes |\phi_n\rangle$$

is defined by the equation

$$\langle \alpha_1 \cdots \alpha_n \mid \psi\rangle = \langle \alpha_1 \mid \phi_1\rangle \cdots \langle \alpha_n \mid \phi_n\rangle$$

Equivalently, the tensor product of three or more vectors can be defined recursively:

$$|\phi_1\rangle \otimes \cdots \otimes |\phi_n\rangle = (|\phi_1\rangle \otimes \cdots \otimes |\phi_{n-1}\rangle) \otimes |\phi_n\rangle$$

We can think about the tensor product of n vectors ϕ_1 through ϕ_n as being the tensor product of the first $n-1$ vectors, followed by the tensor product of the result with the n-th vector.

To find the tensor product of the first $n-1$ vectors, we can apply the same recursive definition, and so on until we reach the tensor product of the first two vectors.

Remark The tensor product of three or more vectors is **multilinear**.

2.1.4 Measurements of probabilistic states

Measurements of compound systems work in the same way as measurements of single systems.

Question 2.1.2

Suppose two systems (X, Y) are together in some probabilistic state. What happens when we measure X and do nothing to Y?

Solution.

1. The probability to observe a particular classical state $x \in \Sigma$ when just X is measured is
$$Pr(\mathsf{X} = x) = \sum_{y \in \Gamma} Pr((\mathsf{X}, \mathsf{Y}) = (x, y))$$
2. There still may exist uncertainty about the classical state of Y, depending on the outcome of the measurement:
$$Pr(\mathsf{Y} = y \mid \mathsf{X} = x) = \frac{Pr((\mathsf{X}, \mathsf{Y}) = (x, y))}{Pr(\mathsf{X} = x)}$$

These formulas can be expressed using the Dirac notation as follows.

Suppose that (X, Y) is in some arbitrary probabilistic state. We start by writing down the probabilistic (X, Y) as a vector:
$$\sum_{(x,y) \in \Sigma \times \Gamma} p_{xy} \, |xy\rangle$$
The entries are denoted by p_{xy} for $x \in \Sigma$ and $y \in \Gamma$.

We can then express $|xy\rangle$ as a tensor product:
$$\sum_{(x,y) \in \Sigma \times \Gamma} p_{xy} \, |xy\rangle = |x\rangle \otimes |y\rangle$$
We can then use the biliearity of the tensor product, specifically the linearity in the second argument, to write this as:
$$\sum_{x \in \Sigma} |x\rangle \otimes \left(\sum_{y \in \Gamma} p_{xy} \, |y\rangle \right)$$
What is happening here is that we are isolating the standard basis vectors on the left hand side, and then grouping the remaining terms on the right hand side. This allows us to see what happens when we measure X and do nothing to Y.

1. The probability that a measurement of X yields an outcome $x \in \Sigma$ is:
$$Pr(\mathsf{X} = x) = \sum_{y \in \Gamma} p_{xy}$$
The probability to get a particular outcome x is the sum of the probabilities of all the outcomes (X, Y) for $y \in \Gamma$.
2. Conditioned on the outcome $x \in \Sigma$, the probabilistic state of Y becomes:
$$\frac{\sum_{y \in \Gamma} p_{xy} \, |y\rangle}{\sum_{c \in \Gamma} p_{xc}}$$
The probability vector that describes the uncertainty of Y is the vector that was in the parentheses on the right hand side, except that it has to be normalized for it to be a probability vector.

Example 2.1.9

Suppose (X, Y) is in the probabilistic state

$$\frac{1}{12}|00\rangle + \frac{1}{4}|01\rangle + \frac{1}{3}|10\rangle + \frac{1}{3}|11\rangle$$

We write this vector as follows:

$$|0\rangle \otimes \left(\frac{1}{12}|0\rangle + \frac{1}{4}|1\rangle\right) + |1\rangle \otimes \left(\frac{1}{3}|0\rangle + \frac{1}{3}|1\rangle\right)$$

Case 1: the measurement of X yields the outcome 0.

$$Pr(\text{outcome is } 0) = \frac{1}{12} + \frac{1}{4} = \frac{1}{3}$$

Conditioned on this outcome, the probabilistic state of Y becomes

$$\frac{\frac{1}{12}|0\rangle + \frac{1}{4}|1\rangle}{\frac{1}{3}} = \frac{1}{4}|0\rangle + \frac{2}{3}|1\rangle$$

Case 2: the measurement of X yields the outcome 1.

$$Pr(\text{outcome is } 1) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

Conditioned on this outcome, the probabilistic state of Y becomes

$$\frac{\frac{1}{3}|0\rangle + \frac{1}{3}|1\rangle}{\frac{2}{3}} = \frac{1}{2}|0\rangle + \frac{1}{2}|1\rangle$$

2.1.5 Operations on probabilistic states

Probabilistic operations on compound systems are represented by stochastic matrices having rows and columns that correspond to the Cartesian product of the individual systems' classical state sets.

Example 2.1.10 (CNOT operation)

A controlled-NOT (CNOT) operation on two bits X and Y:

If X = 1, then perform a NOT operation on Y, otherwise do nothing.

X is the **control bit** that determines whether a NOT operation is applied to the **target bit** Y.

Action on standard basis:
$$\begin{aligned} |00\rangle &\mapsto |00\rangle \\ |01\rangle &\mapsto |01\rangle \\ |10\rangle &\mapsto |11\rangle \\ |11\rangle &\mapsto |10\rangle \end{aligned}$$

Matrix representation:
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Example 2.1.11

Here is a different operation on two bits (X,Y):

With probability 1/2, set Y equal to X, otherwise set X equal to Y.

The matrix representation of this operation is as follows:

$$\begin{pmatrix} 1 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

Question 2.1.3

Suppose we have two probabilistic operations, each on its own system, described by stochastic matrices:

1. $\mathcal{M}$ is an operation on X.
2. $\mathcal{N}$ is an operation on Y.

If we *simultaneously* perform the two operations, how do we describe the effect on the compound system (X,Y)?

Hint: if we simultaneously perform the two operations on their own respective systems, then we are in fact performing an operation on the compound system.

Solution.
placeholder

2.1.6 Tensor products of matrices

The **tensor product** of two matrices, which is analogous to the tensor product of vectors

$$\mathcal{M} = \sum_{\mathsf{a},\mathsf{b}\in\Sigma} \alpha_{\mathsf{ab}} \left|\mathsf{a}\right\rangle\left\langle\mathsf{b}\right| \quad \text{and} \quad \mathcal{N} = \sum_{\mathsf{c},\mathsf{d}\in\Gamma} \beta_{\mathsf{cd}} \left|\mathsf{c}\right\rangle\left\langle\mathsf{d}\right|$$

is the matrix

$$\mathcal{M}\otimes\mathcal{N} = \sum_{\mathsf{a},\mathsf{b}\in\Sigma}\sum_{\mathsf{c},\mathsf{d}\in\Gamma} \alpha_{\mathsf{ab}}\beta_{\mathsf{cd}} \left|ac\right\rangle\left\langle bd\right|$$

An alternative, but equivalent, way to define $\mathcal{M}\otimes\mathcal{N}$ is that it is the unique matrix that satisfies the equation

$$(\mathcal{M}\otimes\mathcal{N})\left|\phi\otimes\psi\right\rangle = \mathcal{M}\left|\phi\right\rangle \otimes \mathcal{N}\left|\psi\right\rangle$$

for every choice of vectors $|\phi\rangle$ and $|\psi\rangle$.

Remark One way to think about this is that the tensor product of matrices works perfectly with the tensor product of vectors. This is, of course, an equation that you would hope is true.

We can define the tensor product in terms of matrices written out:

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1m} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mm} \end{pmatrix} \otimes \begin{pmatrix} \beta_{11} & \cdots & \beta_{1k} \\ \vdots & \ddots & \vdots \\ \beta_{k1} & \cdots & \beta_{kk} \end{pmatrix} = \begin{pmatrix} \alpha_{11}\beta_{11} & \cdots & \alpha_{11}\beta_{1k} & & \alpha_{1m}\beta_{11} & \cdots & \alpha_{1m}\beta_{1k} \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ \alpha_{11}\beta_{k1} & \cdots & \alpha_{11}\beta_{kk} & & \alpha_{1m}\beta_{k1} & \cdots & \alpha_{1m}\beta_{kk} \\ & \vdots & & \ddots & & \vdots & \\ \alpha_{m1}\beta_{11} & \cdots & \alpha_{m1}\beta_{1k} & & \alpha_{mm}\beta_{11} & \cdots & \alpha_{mm}\beta_{1k} \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ \alpha_{m1}\beta_{k1} & \cdots & \alpha_{m1}\beta_{kk} & & \alpha_{mm}\beta_{k1} & \cdots & \alpha_{mm}\beta_{kk} \end{pmatrix}$$